Atmospheric
Chemistry
and Physics

Constraining chemical transport PM$_{2.5}$ modeling outputs using surface monitor measurements and satellite retrievals: application over the San Joaquin Valley

Mariel D. Friberg[1,2], **Ralph A. Kahn**[1], **James A. Limbacher**[1,3], **K. Wyat Appel**[4], **and James A. Mulholland**[2]

[1]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[2]School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[3]Science Systems and Applications Inc., Lanham, MD 20706, USA
[4]US EPA, Research Triangle Park, NC 27711, USA

Correspondence: Mariel D. Friberg (mariel.d.friberg@nasa.gov)

Abstract. Advances in satellite retrieval of aerosol type can improve the accuracy of near-surface air quality characterization by providing broad regional context and decreasing metric uncertainties and errors. The frequent, spatially extensive and radiometrically consistent instantaneous constraints can be especially useful in areas away from ground monitors and progressively downwind of emission sources. We present a physical approach to constraining regional-scale estimates of PM$_{2.5}$, its major chemical component species estimates, and related uncertainty estimates of chemical transport model (CTM; e.g., the Community Multi-scale Air Quality Model) outputs. This approach uses ground-based monitors where available, combined with aerosol optical depth and qualitative constraints on aerosol size, shape, and light-absorption properties from the Multi-angle Imaging SpectroRadiometer (MISR) on the NASA Earth Observing System's Terra satellite. The CTM complements these data by providing complete spatial and temporal coverage. Unlike widely used approaches that train statistical regression models, the technique developed here leverages CTM physical constraints such as the conservation of aerosol mass and meteorological consistency, independent of observations. The CTM also aids in identifying relationships between observed species concentrations and emission sources.

Aerosol air mass types over populated regions of central California are characterized using satellite data acquired during the 2013 San Joaquin field deployment of the NASA Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) project. We investigate the optimal application of incorporating 275 m horizontal-resolution aerosol air-mass-type maps and total-column aerosol optical depth from the MISR Research Aerosol retrieval algorithm (RA) into regional-scale CTM output. The impact on surface PM$_{2.5}$ fields progressively downwind of large single sources is evaluated using contemporaneous surface observations. Spatiotemporal R^2 and RMSE values for the model, constrained by both satellite and surface monitor measurements based on 10-fold cross-validation, are 0.79 and 0.33 for PM$_{2.5}$, 0.88 and 0.65 for NO$_3^-$, 0.78 and 0.23 for SO$_4^{2-}$, 1.00 and 1.01 for NH$_4^+$, 0.73 and 0.23 for OC, and 0.31 and 0.65 for EC, respectively. Regional cross-validation temporal and spatiotemporal R^2 results for the satellite-based PM$_{2.5}$ improve by 30 % and 13 %, respectively, in comparison to unconstrained CTM simulations and provide finer spatial resolution. SO$_4^{2-}$ cross-validation values showed the largest spatial and spatiotemporal R^2 improvement, with a 43 % increase. Assessing this physical technique in a well-instrumented region opens the possibility of applying it globally, especially over areas where surface air quality measurements are scarce or entirely absent.

1 Introduction

To investigate air pollution health effects on humans, population-based epidemiologic time-series studies often use exposure measures derived from regulatory monitoring networks (Laden et al., 2006; Pope et al., 2009; Özkaynak et al., 2009). Even for the continental US, many ambient, ground-level fine particulate matter (PM$_{2.5}$) chemical datasets are acquired only once every 3 or 6 days, and data records at many sites are less than a decade or two long. In addition, the monitors tend to be concentrated in a small number of populated counties, with the exception of the Interagency Monitoring of Projected Visual Environment (IMPROVE) program sites located primarily in US national parks (Hand et al., 2011). Prior to 2009, instrument types and sensitivities varied from monitor to monitor and among monitoring networks (Chow et al., 2010), making comparisons and uncertainty assessment difficult.

Urban-level epidemiological time-series studies often span large geographic regions (Goldstein and Landovitz, 1977; Wade et al., 2006). Especially for long-term exposure analysis, broad regions within or downwind of urban and industrial centers are also of concern due to the presence of distributed populations and natural and agricultural ecosystems. Characterizing spatial variability is fundamental to effectively conducting environmental epidemiologic studies and air quality assessments. Reducing exposure-metric error caused by inadequately characterized spatial variability, which is often much larger than instrument error, can substantially reduce bias and improve precision in epidemiologic results (Ito and Thurston, 1995; Pinto et al., 2004; Goldman et al., 2012). This is particularly relevant for regional-scale studies, where urban-to-rural gradients of ambient surface PM$_{2.5}$ and chemical species concentrations are often lacking. CE1

Although chemical transport model (CTM) simulations provide more complete spatial and temporal coverage than surface monitors, they rely on uncertain inputs about pollution source characteristics that can contain significant biases. The accuracy of the simulated fields is also affected by the accuracy of the simulated meteorology, emissions, and the physical and chemical parameterization schemes specified in the model (Cooke et al., 1999; Monks et al., 2009; Tong and Mauzerall, 2006). Errors in these fields can be identified and sometimes quantified by comparison with coincident ground- and aircraft-based observations. Under satisfactory retrieval conditions, satellite-derived aerosol optical depth (AOD), atmospheric scattering, light absorption, and extinction by suspended particles can be leveraged to constrain the columnar CTM simulations in sparsely monitored areas.

Early space-based PM$_{2.5}$ air quality studies directly correlated satellite-derived AOD from the MODerate resolution Imaging Spectroradiometer (MODIS) instruments and ground-level PM$_{2.5}$ concentrations acknowledged but did not account for particle vertical distribution, day-to-day vari-

ations, and/or aerosol speciation (Chu et al., 2003; Wang and Christopher, 2003; Engel-Cox et al., 2004; Chu, 2006; Gupta and Christopher, 2009; Wallace and Kanaroglou, 2007; Schaap et al., 2009; Zhang et al., 2009; Hu and Rao, 2009; Tsai et al., 2011; Hu et al., 2014). This direct total-column AOD-to-surface PM$_{2.5}$ correlation approach works well when the aerosol is almost entirely concentrated in the near-surface boundary layer but suffers when transported aerosol makes a significant contribution to the total-column AOD or when the boundary layer is deep or variable on short timescales, as has been pointed out by Hidy et al. (2009). Other early studies used surface measurements (Al-Saadi et al., 2005) or CTMs (Liu et al., 2004; Koelemeijer et al., 2006; Mathur, 2008; Van Donkelaar et al., 2010, 2013; Drury et al., 2010; Wang et al., 2010; Boys et al., 2014; Ma et al., 2014) to provide some constraint on aerosol vertical distribution but did not account in detail for either spatial or temporal variations in the relationship between total-column AOD amounts and surface PM$_{2.5}$ concentrations and provided very limited or no aerosol type constraints. Work has been done to improve CTM estimates of surface PM$_{2.5}$ by improving the consistency of aerosol optical properties between models and satellite retrieval algorithms, as well as using CTMs to inform satellite-retrieved aerosol types (Van Donkelaar et al., 2013; Wang et al., 2010; Drury et al., 2010; Li et al., 2015). The Van Donkelaar et al. (2010) study used space-based CALIPSO lidar backscatter profiles to validate the GEOS-Chem model vertical distributions globally, aggregated over a 4-year period. Advanced statistical models that use land-use, meteorological, and relative humidity parameters have been applied to increase the accuracy of AOD-to-PM$_{2.5}$ estimates (Kumar et al., 2007; Di Nicolantonio et al., 2009; Lee et al., 2011; Kloog et al., 2012; Hu et al., 2014; Ma et al., 2014, 2016; Song et al., 2014; Lv et al., 2016). Several of these statistical models are location-specific, and most rely on surface-based data training sets to constrain parameters in statistical models, which are then applied elsewhere. Where training data are limited or entirely absent, there is significant uncertainty with this approach.

The first papers to include some space-based aerosol type information along with AOD from satellites for air quality applications used the Multi-angle Imaging SpectroRadiometer (MISR) spherical vs. nonspherical distinctions to separate airborne dust from spherical particles over the continental US and constrained aerosol vertical distribution and speciated the spherical components with an aerosol transport model (Liu et al., 2007a, b). Subsequent work applied MISR aerosol size and shape constraints over the Indian subcontinent and surrounding areas to map seasonal changes in aerosol type (Dey et al., 2012) and combined MISR particle shape and qualitative light-absorption information to make a first effort at mapping aerosol air mass types over an urban area, i.e., Mexico City (Patadia et al., 2013).

In the current study, we introduce and enhance a physical approach that takes advantage of satellite coverage over

regional scales for estimating ambient PM$_{2.5}$ mass and associated chemically speciated concentrations, as needed in air quality applications. The approach uses ground-based PM$_{2.5}$ measurements, where available, to anchor speciated, near-surface CTM aerosol concentrations. To help constrain the model outputs over extended regions, both MISR total-column AOD and qualitative, column-effective aerosol type observations are also applied when retrieval quality is adequate (generally, where mid-visible AOD values exceed 0.15). Specifically, we map the satellite-retrieved constraints on spherical light-absorbing, spherical non-absorbing, and nonspherical particles to the appropriate aerosol chemical species in the CTM, which is substantially different from previous work. Enhanced aerosol type retrievals from the MISR Research Aerosol (MISR-RA) retrieval algorithm (Kahn et al., 2001; Limbacher and Kahn, 2014), at 275 m horizontal resolution, are at the heart of this new approach.

To demonstrate the method, we apply it over a case study area in the San Joaquin Valley of California during the Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ, http://discover-aq.larc.nasa.gov, last access: 7 August 2018) field campaign in this region, on 6 CE2 days when there is good MISR-RA coverage. The results account for spatiotemporal variability in PM$_{2.5}$ and the chemical component concentrations. The accuracy of estimated concentrations and evaluation of the latest MISR-RA ability to typify urban AOD, aerosol mixtures, and aerosol air masses, is examined by comparing the results with speciated ground observations and standard model-fitting statistics. Section 2 describes the datasets involved, Sect. 3 describes the method and technical approach, and Sect. 4 presents results and validation for our test cases. Conclusions, along with a brief discussion of prospects for a wider application of this approach, are given in Sect. 5, and detailed data and ancillary documentation are provided in the Supplement.

2 Study domain and datasets

2.1 Study domain

The San Joaquin Valley (SJV), which comprises the southern two-thirds of California's Central Valley (about $26\,000\,\mathrm{km}^2$), has long suffered from severe air pollution issues and is among the most studied airsheds in the US (Ngo et al., 2010; Chow et al., 2006). The SJV has complex topography and meteorology, particularly in winter, when low planetary boundary layer (PBL) heights and high pollutant mixing ratios create a challenging environment for chemical transport modeling (Hidy et al., 2009; Appel et al., 2017). This region is surrounded by the Sierra Nevada to the east, the Diablo and Temblor ranges to the west, the Tehachapi Mountains to the south, and the Sacramento Valley to the north (Fig. 1). Although primarily a rural area, the eight counties that comprise

the SJV are home to more than 4 million residents. Despite the semiarid climate, the SJV is one of the world's most productive agricultural regions (Schoups et al., 2005). Its airshed frequently experiences high PM$_{2.5}$ concentrations during the winter due to the combination of relatively dry climate, shallow PBL heights, local source emissions, and the surrounding mountain ranges. The region has been in violation of the PM$_{2.5}$ National Ambient Air Quality Standards for PM$_{2.5}$ annual standard since their inception in 1997 and is the largest PM$_{2.5}$ non-attainment area in the continental US (US EPA, 2018).

The study period for this work was selected to coincide with the DISCOVER-AQ field campaign, which ran from 16 January through 8 February 2013. This campaign was a joint collaboration between NASA, NOAA, US EPA, multiple universities, and several local organizations, with the goal of characterizing air quality in urban areas using satellite, aircraft, vertical profiler, and ground-based measurements. Targeting the 2013 DISCOVER-AQ deployment period for this study provides considerable ground- and aircraft-based measurements for aerosols and fine particulate matter, which we apply as model constraints and for evaluation.

We analyze data for 6 days during the DISCOVER-AQ period for which (1) MISR observations were made over the study region, (2) coincident ground and aircraft observations were acquired, including extensive field-campaign data, and (3) the key observational requirements of relatively cloud-free conditions and the presence of aerosols from different sources are met. Of the 6 days for which we have MISR coverage, the mid-visible AOD exceeds 0.15 on 3 days: 20 January and 3 and 5 February. On lower-AOD days, MISR-RA aerosol type information has higher uncertainty for the current application, and thus the analysis of speciated PM$_{2.5}$ focuses on the higher-AOD days. Of the 3 higher-AOD days, 20 January has the least cloud cover, followed by 5 February, so these days are the main focus of detailed analysis. The method developed here can in the future be applied to many other polluted regions of the world where AOD exceeding 0.15 is common, such as south and east Asia, North Africa, and many major metropolitan areas.

The ground-based, aircraft, and simulation data used in this study are described briefly in the rest of this section, along with the MISR-RA retrieval product.

2.2 Ground-based PM mass and speciated measurements

This study focuses on PM$_{2.5}$ mass and the five components that dominate total PM$_{2.5}$ in the SJV: sulfate (SO$_4$), nitrate (NO$_3$), ammonium (NH$_4$), elemental carbon (EC), and organic carbon (OC). Data files of ambient aerosol particulate matter species concentrations for sites within the SJV for January and February 2013 were obtained from two EPA sources: (1) daily averaged PM$_{2.5}$ Federal Reference Method (FRM) and Federal Equivalence Method (FEM)
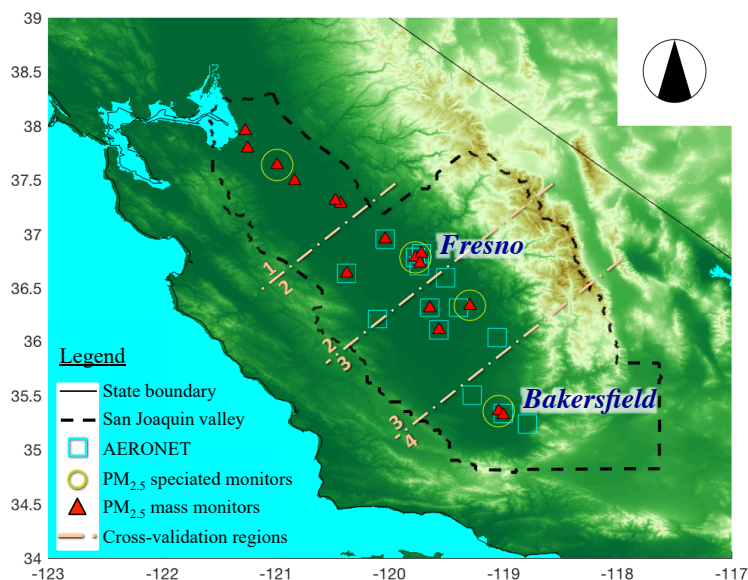
Figure 1. San Joaquin study area shows the ground elevation, EPA AQS and CSN monitors, and AERONET sites during the NASA DISCOVER-AQ flight campaign.

Table 1. EPA AQS and CSN monitor summary statistics for 52 days (6 days).

Pollutant	No. of monitors	Sampling frequency	No. of observations	Mean	SD
PM$_{2.5}$, µg m^{-3}	22 (21)	13 daily; 6 1-in-3; 3 1-in-6	779 (95)	21.20 (28.31)	13.33 (13.51)
PM$_{2.5}$-SO$_4$, µg m^{-3}	7 (6)	6 1-in-3; 1 1-in-6	86 (11)	0.77 (1.13)	0.46 (0.69)
PM$_{2.5}$-NO$_3$, µg m^{-3}	7 (6)	6 1-in-3; 1 1-in-6	86 (11)	7.27 (9.81)	6.11 (7.38)
PM$_{2.5}$-NH$_4$, µg m^{-3}	5 (4)	4 1-in-3; 1 1-in-6	54 (7)	2.07 (3.65)	2.25 (3.32)
PM$_{2.5}$-EC, µg m^{-3}	4 (4)	3 1-in-3; 1 1-in-6	44 (8)	1.28 (1.14)	0.77 (0.34)
PM$_{2.5}$-OC, µg m^{-3}	4 (4)	3 1-in-3; 1 1-in-6	44 (8)	5.25 (5.73)	3.09 (2.48)

mass from the Air Quality System (AQS; https://www.epa.gov/aqs, last access: 7 August 2018) and (2) daily averaged total PM$_{2.5}$ and chemically speciated mass (measurements typically made every third or sixth day) from the Chemical Speciation Network (CSN; Solomon et al., 2014).

FRM compliant data from gravimetric filter-based samplers and FEM compliant data from continuous mass monitors provide spatial variability in PM$_{2.5}$ mass (EPA, 2004). The PM$_{2.5}$ FRM mass is determined gravimetrically by weighing particles on filters pre- and post-deployment. They are equilibrated at a constant relative humidity (30 %–40 %) and temperature (20–23 °C). Monitor locations are shown in Fig. 1, and Table 1 lists monitor summary statistics. Daily PM$_{2.5}$ concentrations measured by the FRM method are considered PM$_{2.5}$ ground truth, i.e., their uncertainties are small compared to those of the other PM$_{2.5}$ values used in this study.

2.3 DISCOVER-AQ AERONET–DRAGON

The AErosol RObotic NETwork (AERONET; Holben et al., 1998) has 10 permanent sun photometer (SP) monitors operating in the study region. During the DISCOVER-AQ mid-January through mid-February 2013 deployment, these monitors were supplemented with an additional 14 temporary monitors termed the Distributed Regional Aerosol Gridded Observation Network (DRAGON) to provide a more regionally dense dataset for satellite validation and in situ comparisons (Fig. 1). AERONET–DRAGON SPs measure AOD in multiple spectral bands from the ultraviolet ($\sim$ 340 nm) to the near-infrared ($\sim$ 1640 nm), with an accuracy within ± 0.015 (Eck et al., 1999).

We use version 2 level 2 (L2) AERONET–DRAGON AOD and Ångström exponent (ANG) data for the 6 study days. The L2 data were sun-calibrated after field deployment, cloud-screened (Smirnov et al., 2000), and quality-controlled. The AOD at 550 nm wavelength is calculated using a quadratic log–log fit to AERONET observations at

shorter and longer wavelengths (Eck et al., 1999). Columnar AODs at 550 nm derived from AERONET are considered to be AOD ground truth in this study.

2.4 Chemical transport model simulations

Simulations of the coupled Weather Research and Forecasting model (WRF; Skamarock et al., 2008), version 3.4, and the Community Multiscale Air Quality model (CMAQ; Byun and Schere, 2006), version 5.0.2, were obtained from the US Environmental Protection Agency (EPA). These hourly atmospheric simulations, at 2 km × 2 km horizontal grid spacing, cover the entire SJV and surrounding major cities during the months of January and February 2013. The CMAQ domain consisted of 35 vertical layers with varying thickness extending from the surface to 50 hPa and an approximately 10 m midpoint for the lowest (surface) model layer. Concentration fields from the fixed 2 km × 2 km horizontal CMAQ grid were downscaled to a horizontal grid of 275 m × 275 m by linear interpolation and used as the reference grid for all subsequent analyses. Emission data were based on the 2011 EPA National Emissions Inventory (EPA, 2015) with 2013 updates to electric generating unit emissions, fire, and mobile sources. Biogenic emissions were generated in-line to CMAQ using the Biogenic Emissions Inventory System (BEIS; http://www.cmascenter.org, last access: 7 August 2018) version 3.14, and the emissions were processed using the Sparse Matrix Operator Kernel Emissions (Houyoux et al., 2000) version 3.5. The carbon bond 2005 chemical mechanism used was CB05TULC (Sarwar et al., 2012; Whitten et al., 2010; Yarwood et al., 2005). The lateral boundary conditions (BCs) for the 2 km simulation were derived from a coupled WRF–CMAQ simulation with 4 km × 4 km horizontal grid spacing, covering the entire state of California and the surrounding areas. Boundary conditions for the 4 km simulation were derived from a 36 km simulation covering the contiguous US, and BCs for the 36 km simulation were provided by a GEOS-Chem simulation (Bey et al., 2001) with the chemical species mapped to the corresponding CMAQ species (Appel et al., 2017).

The EPA conducted a model evaluation of CMAQ v5.0.2 with respect to the scientific updates to v5.1 (Appel et al., 2017). In that study, fine particulate matter simulations were biased low compared to observed concentrations over the SJV during the winter months. Winter PM$_{2.5}$ average mean bias (Model – Observations) in the SJV exceeded $-10 \mu g\, m^{-3}$. Errors in simulated PBL height and mixing were considered to be contributing factors to the January PM$_{2.5}$ underestimation in the SJV. Although CMAQ v5.0.2 is missing several secondary organic aerosol species of anthropogenic volatile organic carbon (i.e., long-chain alkanes and naphthalene) in its aerosol module (AERO6 v5.0.2), the mass contribution of these species to PM$_{2.5}$ during the winter was minimal (less than $\pm 0.5 \mu g\, m^{-3}$) in the SJV (Appel

et al., 2017). At the time this study was conducted, CMAQ v5.1 results were not yet available.

2.5 Satellite observations

The primary satellite resource for this study is the MISR instrument. We supplement the MISR-RA aerosol data with results from the MODIS instruments. They offer more extensive spatial coverage and provide up to two observations per day (one in the morning and one in the early afternoon), though with larger AOD uncertainty over land and with no constraints on aerosol type over land (Levy et al., 2013). We describe these two data sources below.

2.5.1 MISR-RA

MISR was launched along with the first MODIS instrument aboard Terra, the flagship satellite of NASA's Earth Observing System (EOS), in December 1999 (Diner et al., 1998). Since then, Terra has maintained a sun-synchronous orbit, descending from north to south over the equator at a local time of ∼ 10:30. MISR measures upwelling shortwave radiance from Earth at nine distinct view angles along the line of flight (±70.5°, ±60.0°, ±45.6°, ±26.1°, and nadir), in each of the four spectral bands centered at 446, 558, 672, and 866 nm. The one nadir, four forward, and four aft-viewing push broom cameras take approximately 7 min to image a given 380 km wide swath of Earth. Due to swath size, it takes MISR about a week to obtain global coverage. Owing to its multispectral, multiangular capabilities, high spatial resolution (up to 275 m), and highly accurate radiometric calibration (Bruegge et al., 2007; Limbacher and Kahn, 2015, 2017), the MISR-RA is uniquely capable of supporting air quality applications by providing information about aerosol microphysical properties at regional scales. The version 23 4.4 km × 4.4 km MISR Standard Algorithm (MISR-SA) AOD product was not available at the time of the evaluation and is not available at a higher resolution. The MISR-SA has greater inconsistencies in aerosol particle retrievals due to limitations in the aerosol climatology included in the algorithm (74 mixtures for the MISR-SA vs. over 700 for the MISR-RA), poorer surface reflectance assumptions, issues with the radiometric calibration critical for aerosol type retrievals that are corrected in the MISR-RA, details of the acceptance criteria, and the spatial resolution at which the algorithm is run. More details are available in the series of papers by Limbacher and Kahn (2014, 2015, 2017). For particle type retrievals, the MISR-RA performs considerably better than the MISR-SA.

High-resolution (275 m) results from the MISR-RA are used to constrain aerosol concentration and type for the CMAQ model. Because of MISR's ability to sample over a large range of scattering angles (i.e., between about 60 and 160° at midlatitudes), the MISR-RA provides column-averaged information regarding aerosol properties under fa-

vorable retrieval conditions (i.e., cloud-free, low-surface albedo, mid-visible AOD exceeding about 0.15) (Kahn and Gaitley, 2015; Kahn et al., 2010). This information amounts to constraints on particle shape (nonspherical, e.g., dust, vs. spherical AOD fraction), particle size (typically three to five bins, e.g., small, medium, and large AOD fraction, parameterized as the Ångström exponent), and particle light absorption (typically two to four bins, e.g., dirty and clear, represented as single-scattering albedo, SSA = 1.0 − [absorbing AOD] / [total AOD]). Although passive satellite remote sensing can only provide information about aerosol type in two dimensions (column-averaged), a chemical transport model can be used to apportion the amount of aerosol near the surface (e.g., Liu et al., 2007a; van Donkelaar et al., 2010; this study). A brief summary of the MISR-RA retrieval process is provided in Sect. S1 in the Supplement.

Following the work of Patadia et al. (2013), we identify different aerosol air masses by categorizing aerosol based on the qualitative particle size, shape, and light-absorption constraints described above. Specifically, for the purposes of this paper, the 14 aerosol components used by all 774 mixtures included in the refined MISR-RA aerosol climatology (Limbacher and Kahn, 2014) can be organized into three broad aerosol type groups: spherical light-absorbing, spherical non-absorbing, and nonspherical (cirrus is ignored in the current application). Especially at low AOD, the MISR-RA-derived aerosol type sensitivity amounts to no more than these three groupings (Kahn and Gaitley, 2015). However, the general microphysical properties of the three broad aerosol groups (AGs) can be associated with specific chemical species identified in the chemical transport model results, as described below in Sect. 3.2. From the point of view of retrieval sensitivity, these three categories map to common aerosol species as follows (Table S2): (1) light-absorbing carbon (LAC), (2) inorganic ions (II) plus organic matter (OM) plus sea salt (SS), and (3) dust. Section S2 provides a description of how the aggregated AOD retrievals are computed for the spherical absorbing aerosol components and separately for spherical non-absorbing aerosol components. It is well established that MISR AOD retrievals suffer biases for scenes with substantial cloud cover (Witek et al., 2013; Shi et al., 2014; Limbacher and Kahn, 2015). Consistent with both Witek et al. (2013) and Limbacher and Kahn (2015), we present results only for days where clouds cover less than 30 % of the scene within the SJV as indicated by the MISR-RA cloud mask, excluding the rural areas that extend into the Sierra Nevada.

2.5.2 MODIS – MAIAC

To supplement the MISR-RA AOD values where MISR coverage is lacking, we adopt results from the MODIS Multi-Angle Implementation of Atmospheric Correction (MAIAC) advanced algorithm (Lyapustin et al., 2012, 2018), which uses time-series analysis and a combination of pixel- and image-based processing to improve the accuracy of cloud detection, aerosol retrievals, and atmospheric correction for surface retrievals. The following is a brief overview of the MAIAC Collection 6 version 2.0 (C6v2) June 2017 North America release aerosol product. The current study uses the MAIAC Atmospheric Properties Products (MCD19A2), which provide AOD at 0.55 μm. A more detailed description of the MAIAC theoretical background and processing steps can be found in Lyapustin et al. (2012, 2018).

After extensive characterization of the MODIS-observed surface background, the MODIS Level 1B data are gridded to a fixed sinusoidal projection at 1 km horizontal resolution in order to observe the same grid cell over time. Working with a fixed grid not only facilitates the use of polar-orbiting observations as if they were "geostationary", it also simplifies comparison of these datasets to fixed-grid model results and other measurements. In addition to the MODIS instrument on the Terra satellite, a second MODIS flies aboard NASA's Aqua satellite, which crosses the equator on the dayside at 13:30 local time. As a consequence of residual detrending and MODIS Collection 6 (C6) Terra-to-Aqua cross-calibration (Lyapustin et al., 2018), MAIAC currently processes MODIS C6 Terra and Aqua jointly as a single sensor. In addition to considerably greater spatial coverage than MISR, this joint product offers some diurnal spread in sampling relative to the MISR snapshots.

For the time-series analysis, MAIAC utilizes a 4–16-day sliding window technique of scenes from multiple MODIS overpasses to retrieve the surface bidirectional reflectance distribution function (BRDF; 0.466 μm), and spectral regression coefficients (SRCs; 0.466 and 2.13 μm), allowing MAIAC to retrieve AOD at 1 km spatial resolution. Unlike instruments that collect nearly simultaneous observations using push broom scanning, the MAIAC algorithm uses the sliding window technique of consecutive clear MODIS cross-track scanned scenes (i.e., cloud-free conditions with relatively low AOD) over several days to acquire multi-angle sets of observations for each location. This allows MAIAC to retrieve the BRDF from an accumulated, multi-angle set of observations. Working under the assumption that surface reflectance changes rapidly over space and slowly over time (e.g., seasonal changes) helps the MAIAC internal dynamic land–water–snow classification. The algorithm produces well-characterized surface reflectance that improves cloud masking and outperforms traditional pixel-level cloud detection techniques that rely on spatiotemporal analysis (Kloog et al., 2014).

Although AOD is originally retrieved in the MODIS blue band B3 at 0.47 μm, MAIAC offers a standardized and validated AOD product at 0.55 μm. With the exception of smoke and dust aerosol detection, the current algorithm does not retrieve AOD over surfaces occurring at altitudes higher than 3.5 km. Like nearly all satellite-based aerosol retrievals, MAIAC retrievals are unreliable for very low AOD conditions, over mountainous terrain, and over surfaces with very high

albedo. The retrieval conditions that affect this study include low AOD and some cloud-contaminated scenes.

3 Methods

Air quality ground observations are spatially sparse and are often temporally incomplete. CTM simulations provide information that is independent of these observations and are consistent with meteorology and assumed emissions. But they can contain biases and can have difficulty capturing the spatial structure of aerosol dispersion downwind of sources. Satellites offer spatially extensive, mainly column-effective aerosol amount and type that, if included appropriately, can reduce or eliminate fused model–surface measurement biases over large areas, especially regions far from concentrated surface monitors. As there are gaps in the satellite products due to clouds and other retrieval-related issues, we use the model to help complete variable fields at several stages of the process. We also use the model to estimate the near-surface components of column-effective satellite values and use ground-monitor data to constrain and to evaluate the results.

Our approach to fusing surface and satellite-based observations with CMAQ simulations involves five steps, illustrated in Fig. 2. Note that the left side of Fig. 2 tracks the process for deriving total PM$_{2.5}$, whereas the right side presents the flow for speciated PM$_{2.5}$. Blue and orange stars in Fig. 2 indicate where uncertainties are estimated by comparison with AERONET and the EPA ground monitors, respectively. First, the PM$_{2.5}$ mass FRM is reconstructed from the simulated concentration dataset. Then, the CE3 total-column AOD and groupings of model aerosol species that match the spherical light-absorbing, spherical non-absorbing, and nonspherical satellite aerosol type AG are reconstructed from the simulated datasets. In Step 3, spatially complete AOD and grouped AOD maps are produced for each of the 6 study days by combining MISR-RA and MAIAC satellite retrievals with scaled values of the modeled AOD and AG AOD products from Step 2, respectively, to fill any remaining gaps. In Step 3, we reconstruct PM$_{2.5}$ mass FRM from the simulated concentration dataset. Step 4 deconstructs the filled satellite-based total-column AOD and grouped AOD to surface PM$_{2.5}$ and grouped PM$_{2.5}$ mass concentrations using the CTM-speciated vertical distributions, respectively. The fifth and final step involves blending daily averaged ambient ground observations and satellite-based total and grouped PM$_{2.5}$ mass concentrations to estimate daily, spatially refined PM$_{2.5}$ mass and speciated pollutant concentrations.

Overall, the inputs are the speciated ground-monitor data, satellite AOD snapshots and AOD grouped by aerosol type, and the CMAQ model simulations. The outputs are the fused ground-monitor, satellite plus model PM$_{2.5}$ mass concentration field, and speciated versions of this field. A detailed description of the key steps follows.

3.1 Step 1 – CMAQ- and surface-derived PM$_{2.5}$ using reconstruction method

A commonly applied PM$_{2.5}$ mass reconstruction (RM) method, also termed mass closure or material balance, is used to compare the sum of major aerosol components to gravimetrically measured PM$_{2.5}$. This approach also accounts for unmeasured or non-simulated species to avoid double counting. Beginning with Countess et al. (1980), the RM method is used to evaluate measurements, characterize spatiotemporal chemical gradients, estimate source contributions to PM, and calculate visibility impairment due to near-surface aerosol. Additionally, the reconstructed PM$_{2.5}$ mass provides insight into the spatial variations among the speciated data (Frank, 2006; Hand et al., 2011, 2014; Malm et al., 2011). The development of this method, along with the differences between reconstructed and gravimetric mass in the CSN and IMPROVE datasets, have been extensively studied in the US (Malm et al., 2011). Chow et al. (2015) provide a detailed literature review of the various mass reconstruction equations.

For the purposes of this study, the RM equation focuses on the following five representative chemical components, with the relevant references cited: (1) inorganic ions (Chow et al., 1994; Chow and Egami, 1997; Andrews et al., 2000; Nolte et al., 2015); (2) organic matter (DeBell et al., 2006; Hand et al., 2011); (3) EC, also referred to as light-absorbing carbon (Bond and Bergstrom, 2006); (4) crustal material, which includes mineral and soil particles, referred to herein as dust (Malm et al., 1994, 2011); (5) sea salt (Hand et al., 2011); and (6) other elements (Simon et al., 2011), which, in the SJV during the study period, made a negligible contribution to PM$_{2.5}$. The respective references provide details as to how multipliers for each species were derived and summarize the evaluation performed for each major PM component.

In addition to the measured aerosol species of interest, WRF–CMAQ model outputs for relative humidity, temperature, and speciated aerosol vertical distribution were used in the PM$_{2.5}$ mass reconstruction and as needed in the other analysis steps described hereafter. The RM method, excluding negligible "other" elements, was used to compare ground observations, CMAQ results, and satellite-derived concentrations. Table S1 in the Supplement provides a summary of the aerosol equations used for the ground-monitor data and CMAQv5.0.2 simulations. The RM equation used is as fol-
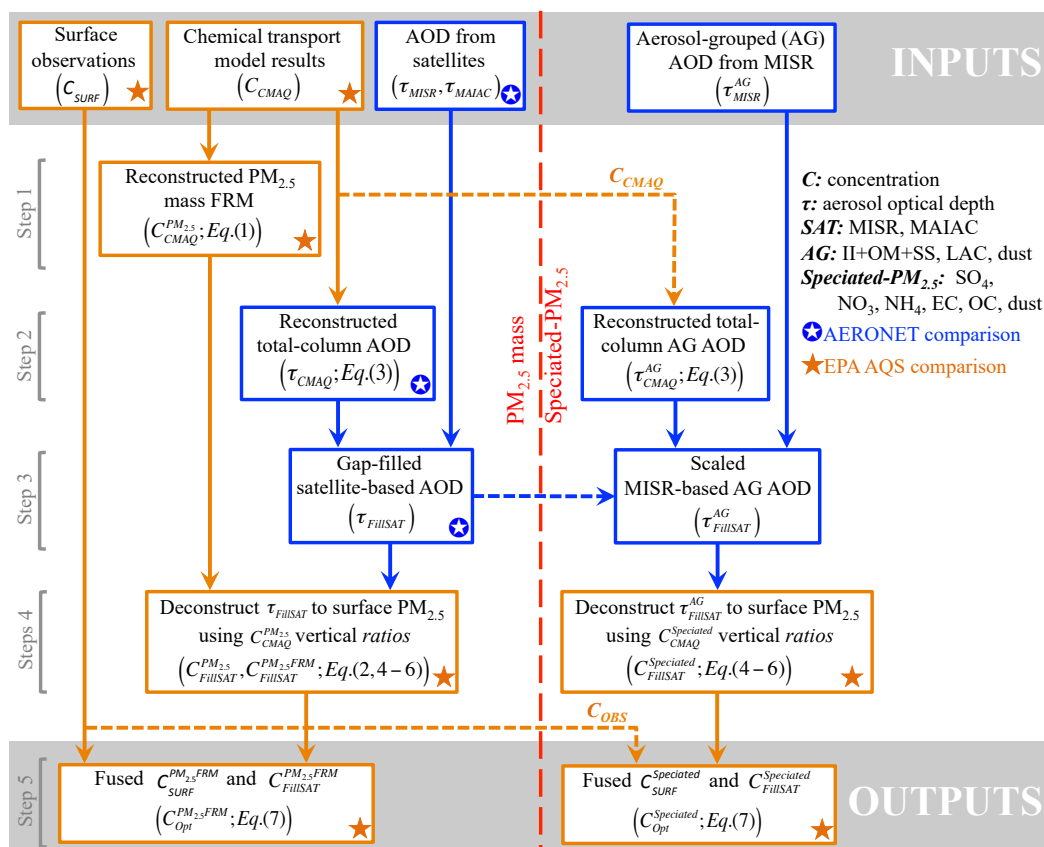
Figure 2. Methods flow chart connecting satellite-retrieved AOD to modeled AOD, PM$_{2.5}$ mass, and PM$_{2.5}$-speciated mass. The parenthetical terms are defined in their respective steps in the methods section.

lows (Eq. A in Chow et al., 2015):

$$RM[\mu g\,m^{-3}] =$$
$$\underbrace{[SO_4^-] + [NH_4^+] + [NO_3^=]}_{\text{inorganic ions}}$$
$$+ \underbrace{1.8[OC]}_{\text{organic matter}} + \underbrace{[EC]}_{\text{light absorbing carbon}} + \underbrace{1.8[Cl^-]}_{\text{sea salt}}$$
$$+ \underbrace{2.2[Al] + 2.49[Si] + 1.63[Ca] + 1.94[Ti] + 2.42[Fe]}_{\text{dust}}. \quad (1)$$

For each of the major chemical components involved, Chow et al. (2015) cover in detail the factors and assumptions required for the RM calculation and those contributing to the comparison with gravimetric mass measurements. These factors include the OM / OC ratio assumptions, carbon sampling and analysis artifacts, ammonium and nitrate volatilization, limitations of using chloride to estimate sea salt content, and water retention by hygroscopic species on filters (Andrews et al., 2000; Rees et al., 2004; Tanner et al., 2004; El-Zanan et al., 2005). Using Eq. (1) to estimate OM from OC for CMAQ output allows for consistency with satellite-derived estimates; in the future, we might expand the

method to include various organic aerosol species explicitly in cases where we have more in situ data.

Following the framework of Eq. (1), the reconstructed PM$_{2.5}$ mass does not account for the positive and negative factors that affect gravimetric and speciated measurements (DeBell et al., 2006; Frank, 2006; Hand et al., 2011; Chow et al., 2015). To close the mass-balance difference between PM$_{2.5}$ FRM gravimetric mass and ambient mass (simulated and measured), the material balance Eq. (1) was adjusted to account for factors affecting gravimetric measurements (Eq. 10 in Frank et al., 2006).

$$PM_{2.5\,FRM}[\mu g\,m^{-3}] = RM - ([NH_4^+]_{loss} + [NO_3^=]_{loss})$$
$$+ [PBW] + [Blank_{FRM}], \quad (2)$$

where ammonium and nitrate volatilization are not captured by gravimetric measurements and thus are accounted as negative artifacts. The particle-bound water (PBW) is the water retained on the filter when particles are sampled and weighed for mass concentration. This concentration is dependent on ionic composition and relative-humidity-dependent species equilibrium prior to laboratory weighing. Blank$_{FRM}$ accounts for the passively collected mass value on "blank" filters. The

limitations and uncertainties of the reconstruction method broken down by major chemical components are discussed in detail elsewhere (Frank, 2006; Chow et al., 2015). The uncertainty estimated for the CMAQ- and satellite-based surface concentrations are discussed in Sect. 4.

3.2 Step 2 – CMAQ-based columnar AOD and AOD subdivided into species-related groups derived using the reconstructed extinction coefficient method

CE4

Section 3.1 summarizes the method applied to calculate the five representative component surface mass concentrations from the surface observations; these components are also used to derive total-column AOD from CMAQ (τ_{CMAQ}). First proposed by Malm et al. (1994), the reconstructed extinction coefficient method was designed to investigate the spatial and temporal variability in haze and visibility impairment in the US as part of IMPROVE. Since then, this method has been continuously upgraded by several researchers (Malm et al., 1994, 2000, 2011; Song et al., 2008; Park et al., 2011). The process estimates total-column extinction AOD using simulated concentrations of II, OM, SS, LAC, and dust (Table S1) assuming externally mixed aerosols with respect to the modeled altitude (z), as follows:

$$\tau = \int \left\{ \underbrace{\sum_i \omega_{(z),i} \beta_{de,i} f_{rh(z),i} C_{(z)i}}_{\text{particle scattering efficiency}} \right.$$
$$\left. + \underbrace{\sum_i (1 - \omega_{(z),i}) \beta_{de,i} f_{rh(z),i} C_{(z)i}}_{\text{particle absorption efficiency}} \right\} dz, \qquad (3)$$

where τ is aerosol extinction optical depth (AOD) at 550 nm, i is the chemical component, ω is the single-scattering albedo (SSA), β_{de} is the specific dry mass extinction efficiency (m^2 g^{-1}), f_{rh} is the hygroscopic growth factors as a function of height, and C is the concentration of chemical component i as a function of height (g m^{-3}).

Equation (3) is further subdivided for dust by size in accordance with the CMAQ Aitken, accumulation, and coarse particles size categories (Park et al., 2011). The empirically based factors and their respective literature sources are summarized in Table S3. The WRF-simulated relative humidity data, rh(z), were used to evaluate the height-dependent hygroscopic growth factors. The ambient particle extinction as a function of height is the sum of the ambient scattering and absorption with respect to altitude (z), which are the two terms in Eq. (3). From Eq. (3), the dimensionless extinction AOD is obtained by multiplying the ambient particle extinction by the vertical atmospheric path height of each CMAQ layer. These are added vertically to obtain columnar AOD values, which are compared to ground- and satellite-based

AOD values in the following subsections to assess uncertainties.

The three CMAQ-based AOD AG (i.e., LAC, II + OM + SS, and dust), indicated in Table S2, are calculated using the five major chemical components derived in Eq. (1). The CMAQ-based total-column AOD AG aggregate is equivalent to the total-column CMAQ-based AOD. The assessment of the uncertainties in these quantities, using a combination of ground-based and satellite total-column measurements, is given in Sect. 3.4 below.

3.3 Step 3 – gap-filled satellite-derived AOD and grouped AOD, using scaled CMAQ-based AOD

To obtain a spatially complete AOD map for each case study day, we combine the MISR-RA-retrieved, MAIAC-retrieved, and CMAQ-based reconstructed AOD products, as CMAQ can simulate values in all grid boxes, regardless of cloud cover, surface brightness, terrain, and aerosol optical thickness. The most relevant factor affecting spatially complete satellite-retrieved AOD in this study is missing retrievals due to the presence of clouds. The combined AOD product is more complete than the MISR-RA or MAIAC product alone.

The Fig. S1 scatterplots show MISR-RA AOD retrievals are higher than those retrieved by MAIAC, and much closer to the AERONET ground-truth values, for the 3 case study days with highest AOD. These scatterplots reinforce the need to scale MAIAC-retrieved AOD before gap-filling MISR-RA-retrieved AOD fields. Based on Fig. S1, a study-specific AOD adjustment was applied to the MAIAC data; in addition, a filter with an upper bound of 0.4 was used for MAIAC retrievals to reduce potential cloud contamination. On days when Aqua and Terra MAIAC C6v2 AOD retrievals on the 1 km fixed sampling grid were available, the MAIAC-Aqua AOD retrievals were used to fill in missing AOD in the MAIAC-Terra AOD maps (as MAIAC-Terra is closest in time to the MISR-RA retrieval) by linearly regressing values from a 15 × 15 MAIAC-Aqua grid cell region centered on the missing MAIAC-Terra cell value. The 1 km gap-filled MAIAC-Terra AOD maps were subsequently downscaled and spatially interpolated (via bilinear interpolation) to match the downscaled CMAQ 275 m × 275 m output grid, referred to herein as gap-filled MAIAC. Before combining retrieved AOD products, the 275 m × 275 m MISR-RA AOD at 558 nm was converted to 550 nm using the retrieved ANG product and the dynamic sampling grid was re-gridded to match the downscaled CMAQ 275 m × 275 m grid. The gap-filled MAIAC product was then used to fill in gaps in the MISR-RA AOD product by linearly regressing values from a 15 × 15 gap-filled MAIAC grid cell region centered on the missing MISR-RA cell value. Larger gaps caused by cloud contamination in the satellite-retrieved AOD were filled using a 7 × 7 grid cell region of CMAQ-reconstructed AOD values, linearly regressed to the satellite-retrieved AOD. This procedure was repeated multiple times as needed until the

satellite retrieval area within the SJV study region was filled, referred herein as τ_{FillSAT}.

A unique component of this work involves the use of the MISR-RA aerosol species-specific groups. Consequently, we produce gap-filled, aerosol-type-grouped AODs from the original MISR-RA-based AG AODs using the AODs from Step 2 CE5 grouped according to the model and following the same gap-filling procedure used for τ_{FillSAT}.

3.4 Uncertainty estimates for model-reconstructed and satellite total-column quantities

Two sets of intermediate analyses are presented where surface-based in situ as well as column-integrated observations are provided as ground truth (i.e., their uncertainties are small compared to those of the other values used in this study). First, satellite-retrieved AOD snapshots are evaluated against coincident AERONET observations. Second, a comparison between daylight-averaged AERONET AOD data, satellite-retrieved AOD snapshots, and model-reconstructed diurnal AOD is presented to determine how well the snapshots represent diurnal values in the study region. This material is presented here rather than in Sect. 4 below because key decisions in the method depend on the results of these comparisons.

3.4.1 Comparison between satellite, CMAQ-reconstructed, and ground-based total-column AOD snapshots at coincident times

Evaluation of MISR-RA (Limbacher and Kahn, 2014) and MAIAC (Lyapustin et al., 2011) AOD has been performed extensively before but not specifically for the study region, where we have considerable ground-truth data. Overall, there were 14 AERONET sites across the SJV (Fig. 1) during the 6 case study days. The number of coincident satellite- and ground-AOD observations is dependent on the swath width of each satellite instrument, the retrieval algorithm used, and the polar-orbiting coverage for a given day. Figure 3 and Table S4 provide scatterplots and a statistical summary, respectively, of AERONET AOD collocated in time and space with the MISR-RA, MAIAC, gap-filled MISR-RA AOD (i.e., τ_{FillSAT}), and CMAQ results. Although AERONET reports AOD at 550 nm, AOD values at 558 nm were calculated for comparison with the MISR-RA AOD retrievals. Only those Terra MAIAC AOD retrievals that were temporally coincident with MISR-RA retrievals were used in this comparison. A window of ± 15 min was applied to select AERONET measurements as spatiotemporally coincident with the satellite overpass, and corresponding CMAQ hourly, reconstructed AOD values were used.

Overall, the MISR-RA AOD compares well with coincident AERONET AOD and tends to outperform MAIAC statistically over the SJV across all our case study days (Table S4). The two best-case days for this analysis are 20 Jan-

uary and 5 February, where AERONET AOD values were relatively high (AOD ≥ 0.15) and there were multiple coincident MISR-RA retrievals across the region. On these days, MAIAC underestimates AOD compared to AERONET, whereas MISR-RA slightly overestimates AOD. Specifically, for 20 January and 5 February, the MISR-RA-to-AERONET AODs had an overall R of 0.91 and 0.99, and an normalized mean absolute error (NME) of 0.08 and 0.12, respectively. For MAIAC, the corresponding values are an overall R of 0.66 and 0.93 and an NME of 0.23 and 0.31, respectively.

The comparison of MISR-RA and MAIAC satellite-retrieved AODs with AERONET also illustrates how gap-filling MISR-RA with scaled and gap-filled MAIAC retrievals produces a more consistent product. For example, the Fig. 3 subplot for 5 February shows that gap-filled MISR-RA (i.e., FillSAT) offers better agreement than gap-filled MAIAC at the averaged AERONET-retrieved AOD value of 0.47. On this specific day and location there is no coincident MISR-RA retrieval, indicating that the gap-filled MISR-RA improvement is due to scaled and gap-filled MAIAC used to gap-fill the MISR-RA AOD snapshot. Further evident from Fig. 3, the CMAQ reconstructed values systematically underestimate AOD relative to AERONET in nearly all cases and exhibit greater scatter, hinting at the possible value of applying the measurements as constraints on the model simulations.

3.4.2 Comparison of satellite-based AOD snapshots with daylight-average ground-based AOD and with daylight- and diurnal-average model-based AOD

Unlike aerosol radiative forcing, which depends on daytime solar heating, conditions during the full diurnal cycle are relevant for many air quality applications. However, AERONET, as well as the satellites, acquire AOD data only during daylight hours, when the sun is well above the horizon. To test the feasibility of using satellite-based AOD snapshot retrievals as proxies for AOD averaged over daylight hours for the study region, we compare the satellite retrievals (MISR-RA, MAIAC, gap-filled MISR-RA) with daylight-averaged AERONET-retrieved AOD results (Fig. S2). We subsequently compare the model daylight- and diurnal-average AODs, as well as the AERONET daylight-average AODs, with the respective short-term values from these data sources (Fig. 4) to assess how well snapshot values represent AOD for entire days in the study region. In places where the snapshots are substantially different from the daylight-average or diurnal-average AOD values, scaled model results would be required to complete the diurnal air quality picture.

For the initial comparison, all retrieved AERONET values per each of the 6 case study days were averaged to obtain a daylight average at each of the 14 sites. For the MISR-RA comparison, we have only the same MISR-RA AOD re-
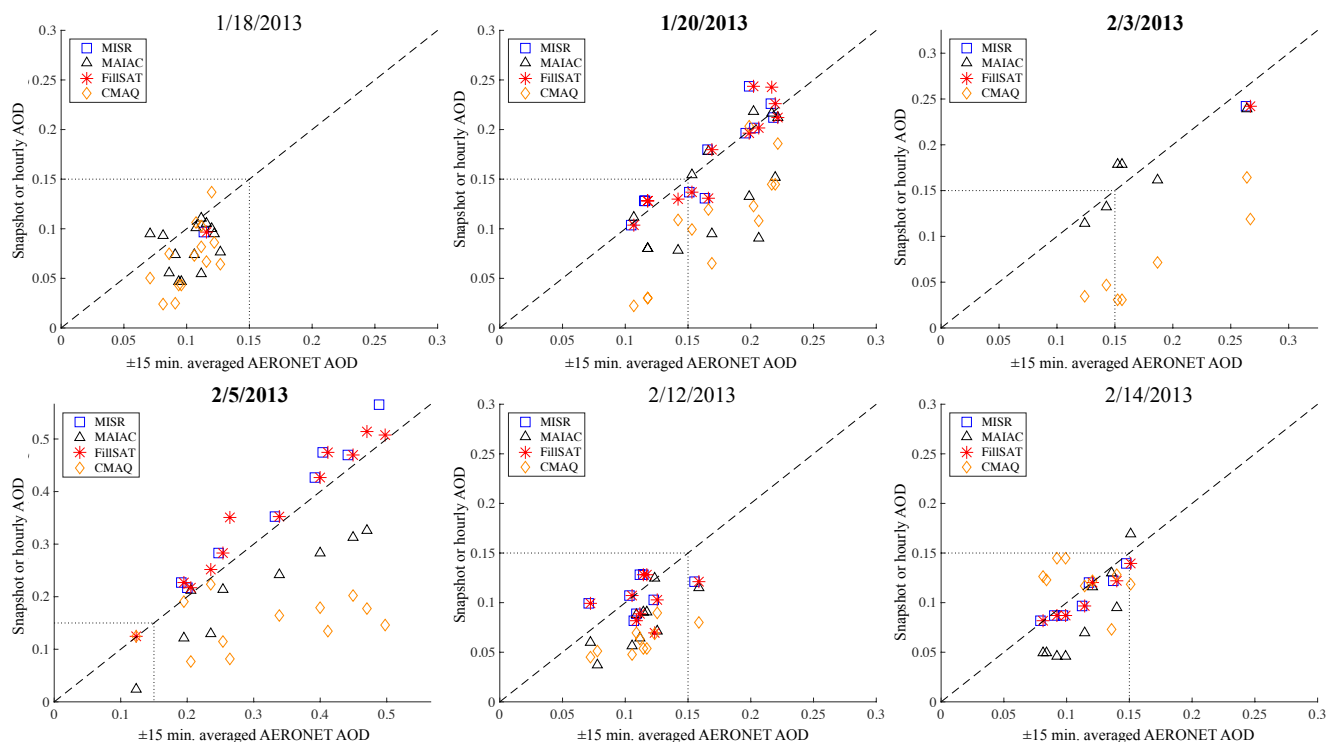
Figure 3. Scatterplot comparisons of AERONET coincidences with MISR-RA, MAIAC, gap-filled MISR-RA, and CMAQ results within ±15 min of Terra overpass time. The MAIAC and AERONET AOD comparisons are plotted at 550 nm, whereas the MISR-RA and AERONET AOD comparisons are at 558 nm; the dotted lines indicate the 0.15 AOD threshold; 1 : 1 dashed lines are shown for reference.

trieval snapshots as in Fig. 3. For the study cases, MAIAC can have multiple Terra and Aqua retrievals over the region during a day, occurring at different times, due to the wide MODIS swath. As such, MAIAC Terra-retrieved AOD "coincident" with MISR-RA overpasses are in some cases gap-filled with other scaled-MAIAC Terra/Aqua retrievals acquired during that day. A third satellite-retrieved AOD product is the gap-filled, primarily MISR-RA-derived AOD (Fill-SAT) described in Sect. 3.3. Also shown in Fig. S2 are the CMAQ reconstructed daylight-average AODs, described in Sect. 3.2.

Overall, the MISR-RA and FillSAT values are very nearly identical, and they tend to serve as better proxies for the daylight-average AERONET values than CMAQ for the study cases. Table S5 contains a statistical summary of the scatterplot data. For the 2 best days of 20 January and 5 February, the retrieved AODs for MISR-RA and gap-filled MISR-RA agree better statistically than the other datasets in terms of correlation and error relative to AERONET daylight-average values. Although the retrieved AODs for the MISR-RA and gap-filled MISR-RA slightly outperform MAIAC for the specific case study days, this relationship is likely to change for different domains and time periods. As such, the technique for gap-filling MISR-RA AOD might need to be dynamic in weighting the MAIAC AOD re-

trievals when applied to other regions. For 20 January and 5 February, the gap-filled MISR-RA-to-daylight-average-AERONET AODs had overall R values of 0.81 and 0.78 and an NME of 0.16 and 0.28, respectively. This comparison indicates the satellite-retrieved AOD quantities are in sufficient agreement with daylight-averaged ground truth to serve as proxies for the daylight-averaged values during the study period.

A procedure for fusing CMAQ model simulations with surface-based measurements is described briefly in Sect. S3 in the supplemental material and in detail in Friberg et al. (2016). This procedure was applied to C_{SURF} and C_{CMAQ} (Fig. 2) to produce C_{FCMAQ}, also referred to as FCMAQ. The additional step allows us to assess how the spatially extensive satellite data affect the results compared to the model constrained only by local surface observations.

To estimate how well the AOD snapshots might characterize the diurnal-average AOD, diurnal-to-hourly ratios for CMAQ and FCMAQ are plotted against AERONET-retrieved AODs acquired within 15 min of the satellite overpasses for each case (Fig. 4 and Table S6). AERONET ratios are plotted as well. The diurnal model and daylight AERONET AOD values are divided by AODs at Terra overpass time within the hour and within 15 min for the model and AERONET ratios, respectively. On 18 and 20 January,
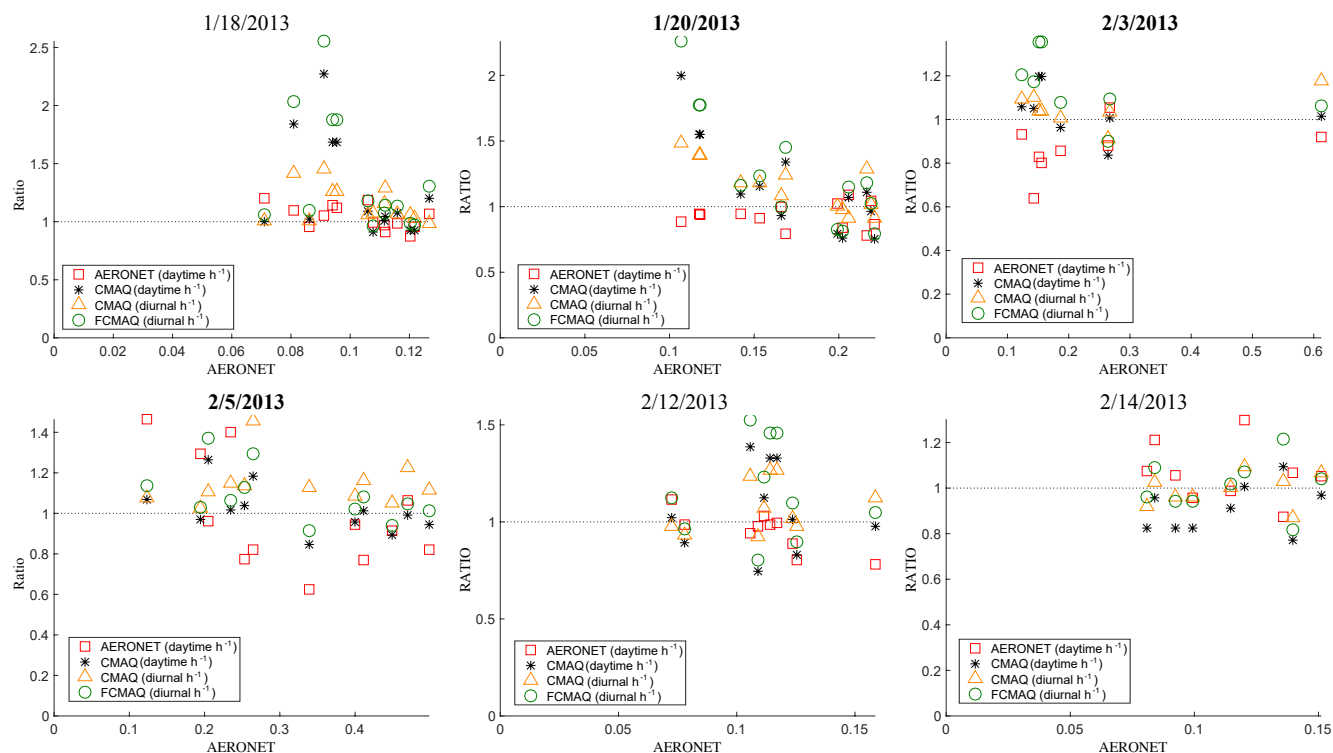
Figure 4. Scatterplots of the ratios of daylight averages to the Terra overpass time vs. AERONET AOD retrievals within ±15 min of Terra overpass time. Two ratios are shown for CMAQ: daytime average-to-hour ratio and diurnal average-to-hour ratio. The FCMAQ ratios shown are the FCMAQ diurnal to CMAQ hour values. The dashed unity lines are included for reference.

FCMAQ and daytime CMAQ ratios exhibit high variability at locations where AERONET ratios were near unity, suggesting that CMAQ diurnal-to-hour ratios are at times spatially biased. But generally, based on model performance, snapshots acquired at Terra overpass time tend to fall within 10 %–20 % of the diurnal-average value, except in some cases when the AOD at overpass time is less than ~ 0.15. At these smaller AODs, a small absolute change in AOD will produce larger percent changes.

One possible reason for the scatter in Fig. 4 is the model representation of transported aerosol. Transported aerosol above the boundary layer is dependent on the lower BCs adopted in the model, and thus it is not always well represented by CMAQ in this region. For example, the model results indicate minimal vertical distribution of dust aerosol, concentrating all the dust within the planetary boundary layer on the study days, whereas transported dust above the boundary layer is likely to be the major nonspherical aerosol species in this region and season (e.g., Liu et al., 2007b). Any biases in dust AOD retrievals are compounded by inaccuracies in the model-based vertical distributions that are applied during the total-column-to-surface decomposition step. The impact of errors in the adopted vertical distribution of aerosols on these results, beyond the scope of the current paper, warrants further investigation. Model aerosol verti-

cal distribution can be further constrained by taking advantage of upwind aerosol elevation retrievals from space-based stereo imaging (MISR), in places where the aerosol sources produce visible plumes and downwind aerosol layer heights from space-based lidar (e.g., CALIPSO) (Kahn et al., 2008).

3.5 Step 4 – deconstructed total-column satellite-measured AOD to surface PM$_{2.5}$ mass and speciated mass concentrations

Using CMAQ-based aerosol vertical profiles, near-surface concentrations ($C_{\text{FillSAT}, z=0}^{\text{PM}_{2.5}\text{FRM}}$ and $C_{\text{FillSAT}, z=0}^{\text{Speciated}}$) are obtained from the total-column satellite AOD (τ_{FillSAT}) and aerosol group AOD ($\tau_{\text{FillSAT}}^{\text{AG}}$) by the following three intermediate steps. As in previous work, the key step amounts to using model-derived ratios of total-column to near-surface aerosol distributions to obtain near-surface values constrained by total-column measurements (e.g., Liu et al., 2004; Van Donkelaar et al., 2010).

In Eq. (4), the column-average dry particle concentrations for the three aerosol groups $\left(\overline{C_{\text{FillSAT}}^{\text{AG}}} \right)$ are calculated from the AODs, τ_{FillSAT}, and $\tau_{\text{FillSAT}}^{\text{AG}}$, by reversing the reconstructed extinction process applied to model-only values in Step 2 (Eq. 3). The same height-stratified hygroscopic growth and specific dry scattering or ab-

sorbing efficiency factors from Step 2 are used here for consistency. The column-average satellite-based AG concentrations $\left(\overline{C_{\mathrm{FillSAT}}^{\mathrm{AG}}}\right)$ are further stratified into the five column-average representative PM chemical components $\left(\overline{\mathrm{Columnar}^{\mathrm{CE6}}\,C_{\mathrm{FillSAT}}^{\mathrm{Speciated}}}\right)$, defined in Step 1 according to Eq. (1), using the CMAQ-based species-to-aerosol group partition in Eq. (5). With $\overline{C_{\mathrm{FillSAT}}^{\mathrm{Speciated}}}$ defined, satellite-based total-column$^{\mathrm{CE7}}$ PM$_{2.5}$ $\left(C_{\mathrm{FillSAT}}^{\mathrm{PM}_{2.5}}\right)$ is obtained using a version of Eq. (1). The satellite-derived column-average concentrations are then scaled to surface-level concentrations by relying on the vertical distribution of the CMAQ simulations of each species in Eq. (6). The satellite-based surface-level PM$_{2.5}$ concentrations $\left(C_{\mathrm{FillSAT},\,z=0}^{\mathrm{PM}_{2.5}}\right)$ are adjusted to reflect PM$_{2.5\,\mathrm{FRM}}$ concentrations using Eq. (2). These relationships were defined in terms of daily AOD and species concentrations.

$$\overline{C_{\mathrm{FillSAT}}^{\mathrm{AG}}} = \frac{\tau_{\mathrm{FillSAT}}^{\mathrm{AG}}}{\int \left(\beta_{\mathrm{de},\,i}\, f_{\mathrm{rh}(z),\,i}\right)\mathrm{d}z} \tag{4}$$

$$\overline{C_{\mathrm{FillSAT}}^{\mathrm{Speciated}}} = \overline{C_{\mathrm{FillSAT}}^{\mathrm{AG}}}\left(\overline{C_{\mathrm{CMAQ}}^{\mathrm{Speciated}}}\middle/\overline{C_{\mathrm{CMAQ}}^{\mathrm{AG}}}\right) \tag{5}$$

$$C_{\mathrm{FillSAT},\,z=0}^{\mathrm{Speciated}} = \overline{C_{\mathrm{FillSAT}}^{\mathrm{Speciated}}}\left(C_{\mathrm{CMAQ},\,z=0}^{\mathrm{Speciated}}\middle/\overline{C_{\mathrm{CMAQ}}^{\mathrm{Speciated}}}\right) \tag{6}$$

3.6 Step 5 – optimized PM$_{2.5}$ FRM and speciated concentrations, derived by fusing satellite-constrained values with ground-monitor data

The optimized concentration dataset (C_{OPT}) closely parallels the surface-measurement-constrained CMAQ simulation described in Eq. (S4). The C_{OPT} dataset is derived by constraining the results with the surface monitor data near their locations and weighting the satellite-constrained concentration values progressively more heavily away from available ground monitors (Fig. 5). Using Eq. (7), the six daily C_{CMAQ} fields coincident with the flight campaign span are replaced with the satellite-derived daily C_{FillSAT} fields, as these were the days when retrieval conditions were adequate to use the data for the current application (see Sect. 2.1 above). With only 11.5 % of the C_{CMAQ} fields changing due to contributions from the surface stations, the weighting factors (W; Eq. S5) and average temporal correlations between the simulations and surface observations (R_2; Eq. S7), across all monitors, did not need to be recalculated. Thus, for this study, C_{OPT} diverges from C_{FCMAQ} for 6 days out of the entire
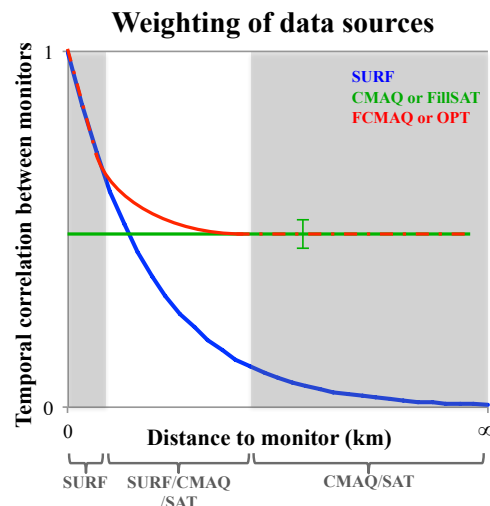
Figure 5. Theoretical plot of fused and optimized dataset weights as a function of distance from surface observations. Three regimes identify the contribution of each dataset towards improving concentration field estimates stratified by distance. As applied here, the exponential decay rate that reflects the temporal Pearson correlation between ground observations as a function of distance is species-specific. The temporal variations between ground observations and CMAQ or FillSAT are more consistent (shown as constant), independent of ground observations, and therefore are not functions of distance to monitor. FCMAQ (surface measurements + model) and OPT (surface + satellite measurements + model) curves show how the strengths of the ground observations and other datasets are maximized using temporal correlations as each grid cell is a function of distance from a ground observation.

study time period.

$$C_{\mathrm{Opt}_{s,t}} = \alpha\,\overline{C_{\mathrm{CMAQ}_s}}^{\beta}\left[W_{s,t}\left\{\frac{C_{\mathrm{SURF}_{sm,t}}}{C_{\mathrm{SURF}_{sm}}}\right\}_{\mathrm{krig}} + \left(1 - W_{s,t}\right)\left\{\frac{C_{\mathrm{FillSAT}_{s,t}}}{\overline{C_{\mathrm{CMAQ}_s}}}\right\}\right] \tag{7}$$

Using the techniques described in the next section, we assess the performance of the optimized surface concentrations in the results section.

3.7 Evaluation of optimized datasets by cross-validation

Three cross-validation techniques are used to evaluate how well the optimized datasets represent diurnal values and to identify biases that arise from different sampling frequencies and the spatial distribution of monitors across the pollutants. First, a 10-fold withholding (10-WH) technique is applied to all species. Then a leave-one-out (LOO) cross-validation method is used for all the species with the exception of PM$_{2.5}$. Finally, a regional holdout (RH) is used only

for PM$_{2.5}$. Brief descriptions of these tests are given here; the results of the tests are discussed in Sect. 4.2 below.

3.7.1 Tenfold cross-validation

The dataset performance was evaluated using a 10-fold cross-validation analysis. For each of 10 independently run trials, a random 10 % of the surface observations were held back per day and each method ("fused", i.e., surface measurements + model, and "optimized", i.e., surface + satellite measurements + model) was applied to simulate the withheld data. The results from the 10 trials were then combined to provide cross-validation results that allow for the exploration of differences in errors based on proximity to monitors. Across monitors and days, the holdout number corresponds to the number of surface observations for each pollutant (Table 1), ranging from 44 for PM$_{2.5}$-OC to 779 for PM$_{2.5}$.

3.7.2 Leave-one-out cross-validation

As an alternative to the 10-WH method, the LOO withholding is applied to the five PM components to overcome the sampling and spatial scarcity. By withholding one location at a time, this location-based cross-validation technique can provide information on how well the CMAQ simulations and satellite-derived concentrations of the fused and optimized datasets, respectively, represent diurnal values at locations further than 50 km from other monitors (see speciated monitor locations in Fig. 1). With some sites containing more than one monitor, collocated monitors were considered to be one location, and thus all monitors at a location were withheld for LOO. This cross-validation technique does not provide much insight when the nearest monitor is in close proximity, as is the case with the PM$_{2.5}$ mass monitors.

3.7.3 Regional holdout cross-validation

A regional withholding technique is used to evaluate fused and optimized PM$_{2.5}$ datasets, as monitor clustering affects the cross-validation results. For each of the cross-validation regions in Fig. 1, all but one of the monitors in a region is withheld, and this is repeated independently for each daily monitor and region. The approach approximates the evaluation of LOO when the distance between monitor locations is large (i.e., > 50 km).

4 Results

Two sets of analyses are presented where surface-based in situ observations are provided as ground truth (i.e., their uncertainties are small compared to those of the other values used in this study). First, modeled and deconstructed satellite-constrained results for PM$_{2.5}$ and PM$_{2.5}$ grouped by species are evaluated against EPA AQS and CSN ground ob-

servations, respectively. For the second set of analyses, cross-validation is used to evaluate satellite-constrained model performance. The main objectives of this section are (1) to evaluate the results of Steps 2–5 as much as possible (for evaluation of Step 1, see Friberg et al., 2017), (2) to assess where, and to what degree, the satellite data help constrain the model PM$_{2.5}$ over an extended region, and (3) where midvisible AOD values exceed 0.15, to also evaluate the satellite-constrained, speciated PM$_{2.5}$.

4.1 Comparison of satellite-constrained and model-based daily PM$_{2.5}$ and speciated component surface concentrations to average daily ground truth

We now compare the model-based (C_{CMAQ}), model-fused-with-ground-monitor (C_{FCMAQ}), deconstructed satellite-constrained (C_{FillSAT}), and optimized (C_{OPT}; model + ground monitor + satellite) daily averaged PM$_{2.5}$ and speciated component concentrations with EPA AQS and CSN observations. Table S7 provides a statistical summary of the comparison between the ground truth and the modeled, fused, satellite-constrained, and optimized results, stratified by pollutant, day, and dataset. Figure 6 presents concentration maps of the four aforementioned datasets with embedded ground-truth PM$_{2.5}$ values and their respective RGB images (depicting cloud cover) for the 3 days with relatively high AOD in the study set (20 January, 3 and 5 February).

Focusing on the area within the SJV, the higher concentration gradients in C_{FillSAT} are due to the application of satellite snapshots. The satellite-constrained concentration snapshots tend to provide more realistic spatial distributions of PM$_{2.5}$ compared to the unconstrained model values, C_{CMAQ}. Specifically, the C_{FillSAT} maps show greater dynamic ranges of values, with localized hotspots over known urban areas, such as Bakersfield (35.4° N, 119.0° W) on 20 January and 5 February and Fresno (36.7° N, 119.8° W) on 3 February. The satellite-constrained snapshot results also tend to agree better with available surface measurements in other high-AOD areas, but cloud contamination and the lack of satellite diurnal sampling affect the C_{FillSAT} values primarily in low-AOD regions. This suggests that the technique will yield increasingly good results when applied in more heavily polluted areas around the globe. Figure S3 presents scatterplots comparing the daily averaged models and the satellite-constrained snapshots of near-surface PM$_{2.5}$ to ground-monitor values. They indicate than diurnal variability is significant in some places and times but not in others. For high-AOD days (20 January, 3 and 5 February), Fig. S3 shows C_{FillSAT} PM$_{2.5}$ is in general agreement with surface observations within the performance range of the model results, and the variability is minimal, especially compared to low AOD days. Of the three relatively high-AOD days, 20 January has the least amount of cloud con-
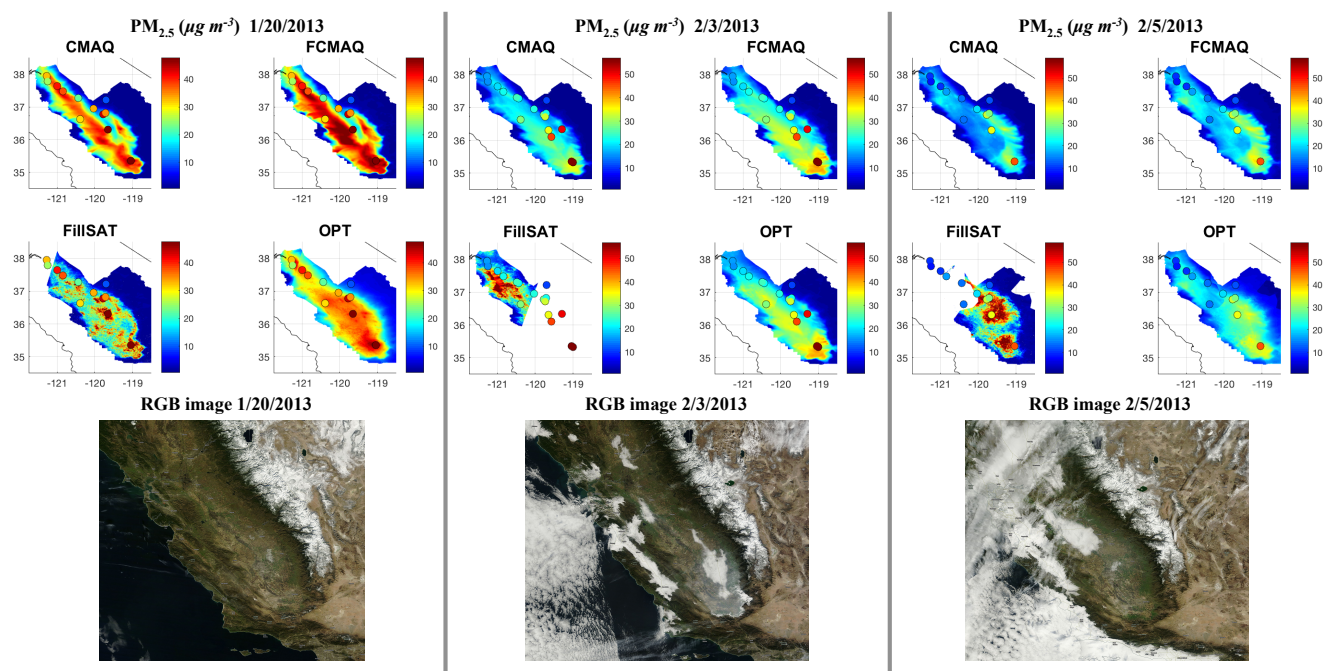
Figure 6. PM$_{2.5}$ FRM calculated concentration maps with monitor observations (filled circles) and RBG images for the 3 days during the study period with highest AOD. The resolution of the concentration maps is 275 m, whereas the size of the observation markers is $\sim 0.1°$ (~ 11.1 km).

tamination, whereas 5 February has the most. Following the Fig. 5 weighting between the datasets, the visible contributions of the C_{CMAQ} and C_{FillSAT} datasets to the C_{FCMAQ} and C_{OPT} PM$_{2.5}$ fields in Fig. 6 occur at distances of a fifth to a half degree (20 to 50 km) beyond a monitor. At or near a ground observation, the C_{OPT} fields are weighted towards the interpolated surface-observation fields, whereas the influence of C_{FillSAT} on C_{OPT} improves the regional behavior and enhances the spatial gradient structure synoptically. For C_{CMAQ} and C_{FillSAT}, the estimated temporal variances are fairly constant and do not depend on distance to the surface observations. The surface observations, rather than model or satellite-based results, dominate the C_{FCMAQ} and C_{OPT} temporal correlations at and near monitor locations, whereas C_{CMAQ} and C_{FillSAT} dominate at distances 20 to 50 km beyond a monitor. As such, the temporal correlations for C_{CMAQ}, C_{FillSAT}, C_{FCMAQ}, and C_{OPT} generally do not approach zero away from the surface stations. For example, on 5 February, the interpolated surface-observation field dominates both the satellite and CMAQ values in the C_{OPT} and C_{FCMAQ} PM$_{2.5}$ maps. The situation at Bakersfield on this day is a bit different. Here the assumed surface monitor uncertainty plays a role, as CMAQ reports a much lower value, the satellite contribution is weighted significantly some distance from the urban center, and the actual difference between the monitor and the C_{OPT} field is about 12.5 %, though the contrast appears large due to the color scale. The satellite contribution is investigated further and quantified in the valida-

tion exercises of the next section, where we systematically decrease the dependence of C_{OPT} fields on surface observations.

Figures 7 and S4 provide speciated NO$_3$, NH$_4$, and SO$_4$ surface concentration maps for 20 January and 3 February, respectively; ground-truth data, available only for 3 February, are included in Fig. 7. For the evaluation of the modeled and satellite-constrained surface concentrations, sparse ground observations of speciated PM have a large impact, especially on the high-AOD days. This is compounded by ground-monitor sampling infrequency, as evident in the correlation ranges (Table S7). Figure S4 demonstrates the ability of satellite aerosol retrievals to characterize the spatial distributions of speciated aerosol air mass types more realistically and consistently than the models across all three species. Unlike for PM$_{2.5}$, there were no speciated monitor measurements available on 20 January, so the OPT results are equal to FillSAT (Fig. S4). Although the C_{CMAQ} and C_{FillSAT} results show agreement around the locations of known emission sources, the satellite-derived aerosol concentrations at the surface show more realistic horizontal dispersion patterns, and the spatial distribution better reflects the likely influence of topographic features. Specifically, during SJV winters, wide horizontal uniformity of ammonium nitrate concentrations is characteristic of this air basin, due to the near-surface inversion (Watson and Chow, 2002). Particulate nitrate is known to form over nonurban areas when high ammonia emissions from the surface and nitric acid, formed
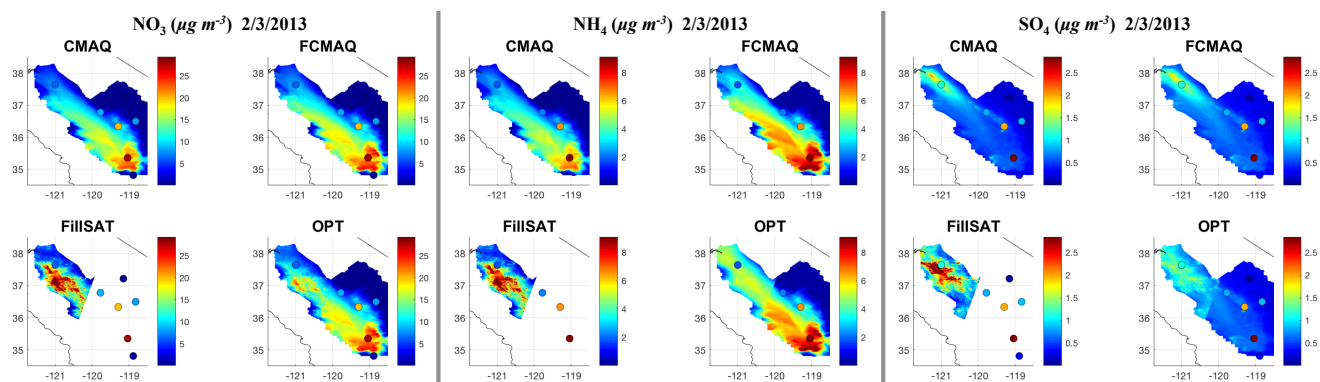
Figure 7. NH$_4$, SO$_4$, and NO$_3$ calculated concentration maps and monitor observations (filled circles) for 3 February.

aloft during nighttime decoupling, mix during the morning collapse of the inversion (Watson and Chow, 2002). Throughout the region, consecutive days with low PBL heights are known to produce increased and spatially more uniform concentrations of fine particulate matter, nitrate, and sulfate (Watson and Chow, 2002). The C_{FillSAT} spatial structure and background concentration ranges of 10–15 µg m^{-3} for nitrate and 4–5 µg m^{-3} for ammonium (Fig. S4) reflect the aforementioned concentration dynamics. The differences between the model and satellite-constrained concentration gradients within the SJV are visible on 20 January and 3 February, and the related surface mixing and plume dispersion are evident, especially in Fig. S4. Given the very limited speciated monitor measurements available, the Fig. S5 scatterplots show C_{FillSAT} provides better agreement than the model and fused-model values.

Comparing the results of the current analysis with previous studies that attempt to apply satellite data to surface air quality assessment is a challenge for the following reasons: (1) limited, nonoverlapping case study domains; (2) disparity in the spatial resolution at which the analyses are performed, which can bias pixel-to-point comparisons; (3) limited number of ground-truth observations; (4) prevalence of statistics that were averaged over entire seasons or years; (5) lack of actual surface concentration statistics reported for the satellite-derived values (i.e., many studies report correlations just between satellite-derived, total-column AOD and surface-based PM$_{2.5}$); and (6) where AOD is the satellite-reported quantity used, algorithm version differences between the AERONET, MISR, and MAIAC products used.

With regard to performance comparisons, the statistical-regression-technique study by Liu et al. (2007b; herein referred to as Liu2007b) is the most similar to the current analysis. Liu2007b compares 54 ground observations to satellite-derived surface concentrations for PM$_{2.5}$ mass and speciated particles over the western US. The statistical regression technique used 3 h averaged CTM (GEOS-Chem) results coincident with Terra overpasses for 2005 at 2° by 2.5° spatial resolution. The Liu2007b regression results with removed

outliers were as follows: PM$_{2.5}$ $R^2 = 0.21$, NO$_3$ $R^2 = 0.23$, SO$_4$ $R^2 = 0.11$, and OC $R^2 = 0.11$. In our study, the spatial R^2 values for PM$_{2.5}$ averaged 0.53 across all days and 0.73 on 20 January, the clearest day with high AOD. The spatial R^2 values for the C_{FillSAT} speciated PM on 12 February, the only day for which we have more than one surface measurement, are 0.48 for NO$_3$, 0.10 for SO$_4$, 0.46 for OC, 0.63 for NH$_4$, and 0.41 for EC.

4.2 Comparison of CMAQ, fused, and optimized datasets to observed concentrations

The model, fused, and optimized datasets are included in the 10-WH cross-validation comparison with the monitor data. The RMSE, mean bias (MB), and the spatiotemporal, temporal, and spatial mean correlations for the five datasets are presented in Table S8. The spatiotemporal $R^2 C_{\text{OPT 10-WH}}$ values are 0.79 for PM$_{2.5}$, 0.88 for NO$_3$, 0.78 for SO$_4$, 1.0 for NH$_4$, 0.73 for OC, and 0.31 for EC. The similarities among the PM$_{2.5}$-speciated component 10-WH cross-validation statistics are affected by low numbers of available observations, sampling frequency, and coincident satellite retrieval data, particularly for NH$_4$ and EC. As a result, when compared to C_{CMAQ}, the $C_{\text{OPT 10-WH}}$ EC results show a 40 % increase in spatial R^2 and 10 % decrease in spatiotemporal R^2, whereas the cross-validation spatiotemporal R^2 values for NH$_4$ are biased high. The SO$_4$ spatial and spatiotemporal R^2 cross-validation results for both C_{FCMAQ} and C_{OPT} show the largest improvement over the unconstrained model, with a 43 % increase compared to the CMAQ simulation performance. The PM$_{2.5}$ temporal and spatiotemporal R^2 cross-validation results are 30 % and 13 % higher than the CMAQ simulations. The C_{OPT} results from the 10-WH cross-validation would normally provide robust cross-validation results that allow for the exploration of error differences based on proximity to monitors. Overall, the statistical improvement between the CMAQ simulations and cross-validated datasets suggest that the empirically based mass reconstruction factors, specific dry efficiencies, and SSA val-

ues adopted were adequate for the SJV domain. The five-city study 10-WH cross-validation spatiotemporal R^2 ranges were 0.81–0.89 for SO$_4$, 0.67–0.83 for PM$_{2.5}$, 0.52–0.72 for NO$_3$, 0.43–0.80 for NH$_4$, and 0.32–0.51 for OC (Friberg et al., 2017). In light of the five-city study, the results for relatively homogeneous pollutants of secondary origin of this study fall within these ranges.

Unlike 10-WH, LOO cross-validation results allow us to leverage the spatial distribution of monitor locations throughout the domain. Table 2 shows the LOO temporal R^2, MB, and RMSE values averaged across monitor locations. The NH$_4$ C_{OPTLOO} results improved the most across the PM$_{2.5}$ component species and outperformed temporal R^2 for C_{FCMAQ} and $C_{FCMAQLOO}$ values by 10 % and 8 %. NH$_4$ cross-validation performance is highest for monitor locations closest to the agricultural emission sources in the southern area of the domain. This finding agrees with the general expectation of aerosol type uncertainties being lowest when the mid-visible AOD is higher than ~ 0.15. For SO$_4$, the cross-validation for both $C_{FCMAQLOO}$ and C_{OPTLOO} datasets shows significant improvements in temporal R^2 and RMSE. For NO$_3$, temporal R^2 of $C_{FCMAQLOO}$ is slightly higher than that of C_{OPTLOO}, whereas the opposite is true for MB. The OC C_{OPTLOO} results are mixed between locations, whereas the EC C_{OPTLOO} shows improvements across all locations.

To explore the PM$_{2.5}$ $C_{FillSAT}$ impact on C_{OPT}, i.e., combining the surface monitor data with the CMAQ simulation plus satellite results, the spatial cross-validation performance assessment of PM$_{2.5}$ C_{OPT} was expanded to include regional holdout (RH), which minimizes the effect of clustered monitors on statistics (Table 3). As expected, removing PM$_{2.5}$-clustered monitors increased the cross-validated dataset reliance of C_{FCMAQ} and C_{OPT} on C_{CMAQ} and $C_{FillSAT}$, thus decreasing temporal R^2 values. PM$_{2.5}$ C_{OPTRH} results are similar for the C_{CMAQ} and $C_{FCMAQRH}$ datasets, with temporal R^2 values of 0.71–0.84 for $C_{FCMAQRH}$ and 0.72–0.83 for C_{OPTRH}. Improvements in the cross-validation results with respect to CMAQ simulations are observed for the northern half of the SJV domain (regions 1 and 2 in Fig. 1). Proximity to emission sources, meteorology, and topography contribute to the performance differences between northern regions 1 and 2 and southern regions 3 and 4. The dominant primary PM$_{2.5}$ mass emission sources (i.e., residential wood combustion and motor vehicles) as well as the major secondary aerosols in the SJV are associated with urban hotspots such as Fresno and Bakersfield (Chen et al., 2007). Winter wind speeds in the SJV are typically below 4 m s^{-1} (Watson and Chow, 2002). As compared to the southern portion of the SJV, the wind speed is slightly higher and is more consistently southeasterly in the northern part of the domain (Cahill et al., 2011; Hayes et al., 1989). During the winter, regional transport occurs when the nocturnal boundary layer is decoupled from the air aloft; as a result, these higher wind speeds aloft tend not to ventilate the surface, intensifying pollutant surface concentrations throughout the SJV (Chow et

Table 2. Comparison of averaged temporal R^2, mean bias, and root means square error values between observations and leave-one-out cross-validation (LOO CV) for 52 days across all locations.

Species	Dataset	Temporal R^2	Mean bias	RMSE
NH4	CMAQ	0.52	0.43	0.94
	FCMAQ	1.00	0.91	1.24
	OPT	1.00	0.70	1.13
	FCMAQ LOO CV	0.56	0.90	1.44
	OPT LOO CV	0.62	0.71	1.39
SO4	CMAQ	0.28	0.02	0.57
	FCMAQ	1.00	0.00	0.12
	OPT	0.99	−0.09	0.11
	FCMAQ LOO CV	0.75	−0.06	0.41
	OPT LOO CV	0.63	−0.13	0.36
NO3	CMAQ	0.73	0.16	0.49
	FCMAQ	1.00	0.26	0.35
	OPT	1.00	0.12	0.31
	FCMAQ LOO CV	0.89	0.14	0.39
	OPT LOO CV	0.85	0.02	0.38
OC	CMAQ	0.68	−0.08	0.36
	FCMAQ	1.00	−0.11	0.14
	OPT	1.00	−0.15	0.13
	FCMAQ LOO CV	0.68	−0.12	0.34
	OPT LOO CV	0.70	−0.14	0.30
EC	CMAQ	0.52	0.31	0.53
	FCMAQ	1.00	0.74	0.85
	OPT	1.00	0.69	0.83
	FCMAQ LOO CV	0.74	0.84	0.87
	OPT LOO CV	0.76	0.80	0.88

al., 1999), whereas dust originating from desert sources to the east and southeast is likely transported aloft.

In summary these results suggest the optimization method is a viable way of constraining CTM simulations using satellite-retrieved information where ground observations are not available, especially where the AOD is higher than in the SJV cases available for the current study. Based on these results, including the satellite data improves short- and long-term spatiotemporal air quality metrics for PM$_{2.5}$ mass and long-term air quality metrics for PM$_{2.5}$-speciated components, especially in areas where surface measurements are lacking.

5 Conclusions

Even in the best-monitored urban areas, ground-based networks have limited spatial coverage. Building on earlier work that produced a method for fusing surface-based measurements with model simulations (Friberg et al., 2016, 2017), the current study relies on both satellite-derived AOD and particle property information contained in the satellite re-

Table 3. Comparison of temporal R^2, mean bias, and root means square error PM$_{2.5}$ values between observations and all simulation, including regional holdout cross-validation (RH CV) for 52 days.

PM$_{2.5}$	Dataset	Temporal R^2	Mean bias	RMSE
	CMAQ	0.68	0.17	0.40
	FCMAQ	1.00	0.10	0.15
Region 1	OPT	1.00	0.09	0.15
	FCMAQ RH CV	0.71	−0.10	0.46
	OPT RH CV	0.73	−0.12	0.46
	CMAQ	0.63	−0.04	0.33
	FCMAQ	0.99	0.05	0.18
Region 2	OPT	0.99	0.03	0.16
	FCMAQ RH CV	0.75	0.05	0.33
	OPT RH CV	0.72	0.03	0.36
	CMAQ	0.77	−0.11	0.30
	FCMAQ	1.00	−0.15	0.17
Region 3	OPT	1.00	−0.17	0.17
	FCMAQ RH CV	0.76	0.06	0.31
	OPT RH CV	0.76	0.02	0.32
	CMAQ	0.82	−0.11	0.34
	FCMAQ	1.00	−0.19	0.24
Region 4	OPT	1.00	−0.23	0.23
	FCMAQ RH CV	0.84	−0.07	0.41
	OPT RH CV	0.83	−0.11	0.39

trievals as additional constraints on the model outputs. The strength of the satellite data is broad spatial coverage, providing radiances that tend to have uniform quality over space and time compared to most suborbital observation datasets. The satellite provides vastly more spatial coverage than the surface stations alone, and this is especially important downwind of major pollution sources. The main limitations of the satellite data are a lack of vertical discrimination in most situations, a lack of diurnal coverage, CE8 and only crude aerosol type sensitivity, especially at low AOD. The physical approach presented here uses CTM simulation along with surface-based measurements to address these limitations. Where satellite data are missing or where the AOD is too low to provide reliable aerosol types from the MISR-RA, the method relies on the model, tuned, to the extent possible, by satellite and surface measurements.

Satellite and ground-based aerosol measurements were combined with numerical model simulations to (1) generate aerosol air mass type maps covering the central California test region for the DISCOVER-AQ campaign time period in 2013, (2) explore the viability of using satellite data to improve aerosol air mass type mapping over extended regions, and (3) contribute regional context to what is known about air pollution sources and trends from point sampling monitors.

Satellite data help capture PM$_{2.5}$ distributions over large, under-sampled or un-sampled regions, and its fusion with model results tends to represent spatial gradients better than the unconstrained model. Applied appropriately, satellite data can also improve speciated PM$_{2.5}$ where AOD is sufficiently high (generally mid-visible AOD $> \sim 0.15$ in the study region). We used retrievals from the MISR-RA, to take advantage of the higher spatial resolution and greater aerosol type accuracy and precision compared to the standard products. However, to avoid overinterpreting the data, we classified the satellite aerosol type results into three broad groups for application as a model constraint: spherical light-absorbing, spherical non-absorbing, and nonspherical. The satellite-constrained concentration maps are spatially consistent with topography, typifying localized hotspots over known urban areas and exhibiting realistic dispersion patterns in the SJV. Comparison with daylight-averaged AERONET and diurnally averaged CMAQ modeling demonstrated that, for AOD $> \sim 0.15$ and with outliers removed, the satellite-derived snapshots represent the diurnal values within 10 %–20 % for the study cases. Furthermore, satellite-derived PM$_{2.5}$ is in agreement with surface observations, to within the scatter of unconstrained model results, and variability was reduced on higher AOD days. These results suggest satellite retrievals in conjunction with model results can improve PM$_{2.5}$ spatial characterization in situations where the AOD is sufficiently high. The satellite aerosol retrievals also represent the spatial distributions of speciated aerosol air mass types more realistically and consistently than the unconstrained model and the model constrained only by surface monitor data for nitrate, ammonium, and possibly also sulfate.

For the current study, model-based aerosol vertical distributions were used to address the lack of profile measurements. However, model aerosol vertical distribution could be constrained on large scales with space-based stereo imaging (e.g., from MISR) near emission sources, at least where plumes are visible in the imagery, and with space-based lidar (e.g., CALIPSO) downwind of sources. Diurnal sampling, the second major limitation in the current satellite application, can be assessed and corrected where needed with a model that has been scaled to available satellite snapshots. Comparison of diurnal variation results to other studies was hindered by the following factors: (1) the unique weather pattern and pollution transport characteristic of the SJV (i.e., persistent inversion and very low PBL height), (2) differences in product version uncertainty (i.e., AERONET versions between this and earlier studies), and (3) disparity in satellite-retrieved spatial resolution (i.e., biases in earlier studies due to coarser spatial resolution). Future research assessing diurnal sampling could benefit from the inclusion of Visible Infrared Imaging Radiometer Suite (VIIRS) instrument datasets, such as daylight-retrieved AOD (Jackson et al., 2013) and the day/night band as an estimate of PM$_{2.5}$ surface change (Wang et al., 2016). Eventually, AOD and possibly speciated AOD from geostationary platforms will provide at least daylight if not fully diurnal values.

Under adequate observing conditions, the technique presented here improves the representation of pollutant spatial distributions in air quality models downwind of emission sources. It is physically based in that it leverages components of a CTM such as the meteorology, conservation of aerosol mass, and assumed emissions CE9 and complements statistical approaches that rely on tuning parameters in a regression-type model. The new method offers the ability to compare satellite-derived PM$_{2.5}$ and speciated concentrations directly to surface measurements. Although the study domain and time frame did not offer the high AOD levels where this method would work best, the SJV offered a substantial quantity of suborbital observations for assessing the results, due to the DISCOVER-AQ campaign.

Expanding this work by applying the technique to the other areas with key ground measurements (i.e., Baltimore DISCOVER-AQ campaign) is a possible next step toward establishing the strengths and limitations of the method. The technique takes advantage of the stable (i.e., consistent), long-term satellite observations that offer global coverage and provides speciated constraints based on retrieved microphysical properties for AOD retrievals above about 0.15. Once the aforementioned analyses are completed, the method will likely be applied to a selection of globally distributed urban regions that are downwind of sources, in locations where particulate pollution levels tend to be high.

Data availability. AQS data are available on the EPA website (https://www.epa.gov/aqs TS1 ; last access: 7 August 2018). CMAQ simulations were obtained by request from the EPA. AERONET/DRAGON data are available on the NASA website (ftp://aeronet.gsfc.nasa.gov; last access: 7 August 2018). The DISCOVER-AQ field data are available on the NASA website (https://doi.org/10.5067/Aircraft/DISCOVER-AQ/Aerosol-TraceGas; last access: 7 August 2018). The MAIAC satellite data are available on the NASA website (https://ladsweb.modaps.eosdis.nasa.gov; last access: 7 August 2018). The MISR satellite data were obtained by request from NASA.

The Supplement related to this article is available online at https://doi.org/10.5194/acp-18-1-2018-supplement.

Author contributions. MDF and RAK designed the study. KWA provided CMAQ data, interpretation, and/or CE10 analysis. RAK and JAL developed and provided MISR-RA code. MDF interpreted and analyzed the satellite data with help from RAK and JAL. MDF prepared the paper with contributions from all coauthors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the CSRA for creating the WRF meteorological inputs and emissions data used in the various model simulations; the US EPA for creating the CMAQ simulations; the US EPA for establishing and maintaining the AQS and CSN sites used in this work and the AQS datasets; the DISCOVER-AQ project for providing some of the field data used in this work; and the NASA AERONET network and its principal investigators, as well as their staff, for establishing and maintaining the AERONET sites used in this work. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. The work of Mariel D. Friberg is supported under grant number NNX13AR89H issued through the NASA Education Minority University Research Education Project (MUREP) as part of the NASA Harriett G. Jenkins Graduate Fellowship program. The work of Ralph A. Kahn is supported in part by NASA's Climate and Radiation Research and Analysis Program under Hal Maring and NASA's Atmospheric Composition Program under R. Eckman TS2 . Last but not least, we thank our anonymous reviewers, who provided comments that have helped us refine this paper.

Edited by: Qiang Zhang
Reviewed by: two anonymous referees

References

Al-Saadi, J., Szykman, J., Pierce, R. B., Kittaka, C., Neil, D., Chu, D. A., Remer, L., Gumley, L., Prins, E., and Weinstock, L.: Improving national air quality forecasts with satellite aerosol observations, B. Am. Meteorol. Soc., 86, 1249–1261, 2005.

Andrews, E., Saxena, P., Musarra, S., Hildemann, L. M., Koutrakis, P., McMurry, P. H., Olmez, I., and White, W. H.: Concentration and composition of atmospheric aerosols from the 1995 SEAVS experiment and a review of the closure between chemical and gravimetric measurements, J. Air Waste Manage., 50, 648–664, 2000.

Appel, K. W., Napelenok, S. L., Foley, K. M., Pye, H. O. T., Hogrefe, C., Luecken, D. J., Bash, J. O., Roselle, S. J., Pleim, J. E., Foroutan, H., Hutzell, W. T., Pouliot, G. A., Sarwar, G., Fahey, K. M., Gantt, B., Gilliam, R. C., Heath, N. K., Kang, D., Mathur, R., Schwede, D. B., Spero, T. L., Wong, D. C., and Young, J. O.: Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1, Geosci. Model Dev., 10, 1703–1732, https://doi.org/10.5194/gmd-10-1703-2017, 2017.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res.-Atmos., 106, 23073–23095, 2001.

Bond, T. C. and Bergstrom, R. W.: Light absorption by carbonaceous particles: An investigative review, Aerosol. Sci. Tech., 40, 27–67, 2006.

Boys, B., Martin, R., Van Donkelaar, A., MacDonell, R., Hsu, N., Cooper, M., Yantosca, R., Lu, Z., Streets, D., and Zhang, Q.: Fifteen-year global time series of satellite-derived fine particulate matter, Environ. Sci. Technol., 48, 11109–11118, 2014.

Please note the remarks at the end of the manuscript.

Bruegge, C. J., Diner, D. J., Kahn, R. A., Chrien, N., Helmlinger, M. C., Gaitley, B. J., and Abdou, W. A.: The MISR radiometric calibration process, Remote Sens. Environ., 107, 2–11, 2007.

Byun, D. and Schere, K. L.: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system, Appl. Mech. Rev., 59, 51–77, 2006.

Cahill, T. A., Barnes, D. E., Spada, N. J., Lawton, J. A., and Cahill, T. M.: Very fine and ultrafine metals and ischemic heart disease in the California central valley 1: 2003–2007, Aerosol Sci. Tech., 45, 1123–1134, 2011.

Chen, L.-W. A., Watson, J. G., Chow, J. C., and Magliano, K. L.: Quantifying PM$_{2.5}$ source contributions for the San Joaquin Valley with multivariate receptor models, Environ. Sci. Technol., 41, 2818–2826, 2007.

Chow, J. C. and Egami, R. T.: San Joaquin Valley 1995 integrated monitoring study: Documentation, evaluation, and descriptive data analysis of PM$_{10}$, PM$_{2.5}$, and precursor gas measurements, Prepared for the California Regional Particulate Air Quality Study, California Air Resources Board, Sacramento, CA, Desert Research Institute, Reno, NV, USA, 1997.

Chow, J. C., Watson, J. G., Fujita, E. M., Lu, Z., Lawson, D. R., and Ashbaugh, L. L.: Temporal and spatial variations of PM$_{2.5}$ and PM$_{10}$ aerosol in the Southern California air quality study, Atmos. Environ., 28, 2061–2080, 1994.

Chow, J. C., Watson, J. G., Lowenthal, D. H., Hackney, R., Magliano, K., Lehrman, D., and Smith, T.: Temporal Variations of PM$_{2.5}$, PM$_{10}$, and Gaseous Precursors during the 1995 Integrated Monitoring Study in Central California, J. Air Waste Manage., 49, 16–24, 1999.

Chow, J. C., Chen, L. W. A., Watson, J. G., Lowenthal, D. H., Magliano, K. A., Turkiewicz, K., and Lehrman, D. E.: PM$_{2.5}$ chemical composition and spatiotemporal variability during the California Regional PM$_{10}$/PM$_{2.5}$ Air Quality Study (CRPAQS), J. Geophys. Res., 111, D10S04, https://doi.org/10.1029/2005JD006457, 2006.

Chow, J. C., Watson, J. G., Chen, L.-W. A., Rice, J., and Frank, N. H.: Quantification of PM$_{2.5}$ organic carbon sampling artifacts in US networks, Atmos. Chem. Phys., 10, 5223–5239, https://doi.org/10.5194/acp-10-5223-2010, 2010.

Chow, J. C., Lowenthal, D. H., Chen, L. W. A., Wang, X., and Watson, J. G.: Mass reconstruction methods for PM$_{2.5}$: a review, Air Qual. Atmos. Hlth., 8, 243–263, 2015.

Chu, D. A.: Analysis of the relationship between MODIS aerosol optical depth and PM$_{2.5}$ in the summertime U.S., Proc. SPIE 6299, Remote Sensing of Aerosol and Chemical Gases, Model Simulation/Assimilation, and Applications to Air Quality, 629903, https://doi.org/10.1117/12.678841, 2006.

Chu, D. A., Kaufman, Y., Zibordi, G., Chern, J., Mao, J., Li, C., and Holben, B.: Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS), J. Geophys. Res., 108, 4661, https://doi.org/10.1029/2002JD003179, 2003.

Cooke, W., Liousse, C., Cachier, H., and Feichter, J.: Construction of a 1 × 1 fossil fuel emission data set for carbonaceous aerosol and implementation and radiative impact in the ECHAM4 model, J. Geophys. Res.-Atmos., 104, 22137–22162, 1999.

Countess, R. J., Wolff, G. T., and Cadle, S. H.: The Denver winter aerosol: a comprehensive chemical characterization, Japca. J. Air Waste Ma., 30, 1194–1200, 1980.

DeBell, L. J., Gebhart, K. A., Hand, J. L., Malm, W. C., Pitchford, M. L., Schichtel, B. A., and White, W. H.: Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States: Report IV, Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, 217–218, http://vista.cira.colostate.edu/Improve/wp-content/uploads/2016/04/IMPROVE_Report_IV.pdf (last access: 7 August 2018), 2006.

Dey, S., Di Girolamo, L., van Donkelaar, A., Tripathi, S. N., Gupta, T., and Mohan, M.: Variability of outdoor fine particulate (PM$_{2.5}$) concentration in the Indian subcontinent: a remote sensing approach, Remote Sens. Environ., 127, 153–161, 2012.

Di Nicolantonio, W., Cacciari, A., and Tomasi, C.: Particulate matter at surface: Northern Italy monitoring based on satellite remote sensing, meteorological fields, and in-situ samplings, IEEE J. Sel. Top. Appl., 2, 284–292, 2009.

Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V., Ackerman, T. P., Davies, R., and Gerstl, S. A. W.: Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview, IEEE T. Geosci. Remote, 36, 1072–1087, 1998.

Drury, E., Jacob, D. J., Spurr, R. J., Wang, J., Shinozuka, Y., Anderson, B. E., Clarke, A. D., Dibb, J., McNaughton, C., and Weber, R.: Synthesis of satellite (MODIS), aircraft (ICARTT), and surface (IMPROVE, EPA-AQS, AERONET) aerosol observations over eastern North America to improve MODIS aerosol retrievals and constrain surface aerosol concentrations and sources, J. Geophys. Res., 115, D14204, https://doi.org/10.1029/2009JD012629, 2010.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, J. Geophys. Res.-Atmos., 104, 31333–31349, 1999.

El-Zanan, H. S., Lowenthal, D. H., Zielinska, B., Chow, J. C., and Kumar, N.: Determination of the organic aerosol mass to organic carbon ratio in IMPROVE samples, Chemosphere, 60, 485–496, 2005.

Engel-Cox, J. A., Holloman, C. H., Coutant, B. W., and Hoff, R. M.: Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality, Atmos. Environ., 38, 2495–2509, 2004.

Frank, N. H.: Retained Nitrate, Hydrated Sulfates, and Carbonaceous Mass in Federal Reference Method Fine Particulate Matter for Six Eastern U.S. Cities, J. Air Waste Manage., 56, 500–511, 2006.

Friberg, M. D., Zhai, X., Holmes, H. A., Chang, H. H., Strickland, M. J., Sarnat, S. E., Tolbert, P. E., Russell, A. G., and Mulholland, J. A.: Method for fusing observational data and chemical transport model simulations to estimate spatiotemporally resolved ambient air pollution, Environ. Sci. Technol., 50, 3695–3705, 2016.

Friberg, M. D., Kahn, R. A., Holmes, H. A., Chang, H. H., Sarnat, S. E., Tolbert, P. E., Russell, A. G., and Mulholland, J. A.: Daily ambient air pollution metrics for five cities: Evaluation of data-

fusion-based estimates and uncertainties, Atmos. Environ., 158, 36–50, 2017.

Goldman, G. T., Mulholland, J. A., Russell, A. G., Gass, K., Strickland, M. J., and Tolbert, P. E.: Characterization of ambient air pollution measurement error in a time-series health study using a geostatistical simulation approach, Atmos. Environ., 57, 101–108, 2012.

Goldstein, I. F. and Landovitz, L.: Analysis of air pollution patterns in New York City – I, Can one station represent the large metropolitan area?, Atmos. Environ., 11, 47–52, 1977.

Gupta, P. and Christopher, S. A.: Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach, J. Geophys. Res., 114, D14205, https://doi.org/10.1029/2008JD011496, 2009.

Hand, J. L., Copeland, S. A., Day, D. E., Dillner, A. M., Indresand, H., Malm, W. C., McDade, C. E., Moore Jr., C. T., Pitchford, M. L., and Schichtel, B. A.: IMPROVE (Interagency Monitoring of Protected Visual Environments): Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States, Report V, CIRA Report, Colorado State University, Fort Collins, CO, http://vista.cira.colostate.edu/improve/wp-content/uploads/2016/08/IMPROVE_V_FullReport.pdf (last access: 7 August 2018), 2011.

Hand, J. L., Schichtel, B. A., Malm, W. C., Copeland, S., Molenar, J. V., Frank, N., and Pitchford, M.: Widespread reductions in haze across the United States from the early 1990s through 2011, Atmos. Environ., 94, 671–679, 2014.

Hayes, T. P., Kinney, J. J., and Wheeler, N. J.: California surface wind climatology, California Air Resources Board, Aerometric Data Division, Aerometric Projects and Laboratory Branch, Meteorology Section, Sacramento, CA, 1989.

Hidy, G. M., Brook, J. R., Chow, J. C., Green, M., Husar, R. B., Lee, C., Scheffe, R. D., Swanson, A., and Watson, J. G.: Remote sensing of particulate pollution from space: have we reached the promised land?, J. Air Waste Manage., 59, 1130–1139, 2009.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., and Nakajima, T.: AERONET – A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 1998.

Houyoux, M. R., Vukovich, J. M., Coats, C. J., Wheeler, N. J., and Kasibhatla, P. S.: Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project, J. Geophys. Res.-Atmos., 105, 9079–9090, 2000.

Hu, X., Waller, L. A., Lyapustin, A., Wang, Y., and Liu, Y.: 10-year spatial and temporal trends of PM$_{2.5}$ concentrations in the southeastern US estimated using high-resolution satellite data, Atmos. Chem. Phys., 14, 6301–6314, https://doi.org/10.5194/acp-14-6301-2014, 2014.

Hu, Z. and Rao, K. R.: Particulate air pollution and chronic ischemic heart disease in the eastern United States: a county level ecological study using satellite aerosol data, Environ. Health, 8, 26, https://doi.org/10.1186/1476-069X-8-26, 2009.

Ito, K. and Thurston, G.: Daily PM$_{10}$/mortality associations: an investigations of at-risk subpopulations, J. Expo. Anal. Env. Epid., 6, 79–95, 1995.

Jackson, J. M., Liu, H., Laszlo, I., Kondragunta, S., Remer, L. A., Huang, J., and Huang, H. C.: Suomi-NPP VIIRS aerosol algorithms and data products, J. Geophys. Res.-Atmos., 118, 12673–12689, https://doi.org/10.1002/2013JD020449, 2013.

Kahn, R., Banerjee, P., and McDonald, D.: Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean, J. Geophys. Res., 106, 18219–18238, 2001.

Kahn, R. A. and Gaitley, B. J.: An analysis of global aerosol type as retrieved by MISR, J. Geophys. Res.-Atmos., 120, 4248–4281, 2015.

Kahn, R. A., Chen, Y., Nelson, D. L., Leung, F. Y., Li, Q., Diner, D. J., and Logan, J. A.: Wildfire smoke injection heights: Two perspectives from space, Geophys. Res. Lett., 35, L04809, https://doi.org/10.1029/2007GL032165, 2008.

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A., and Holben, B. N.: Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network, J. Geophys. Res.-Atmos., 115, D23209, https://doi.org/10.1029/2010JD014601, 2010.

Kloog, I., Melly, S. J., Ridgway, W. L., Coull, B. A., and Schwartz, J.: Using new satellite based exposure methods to study the association between pregnancy PM$_{2.5}$ exposure, premature birth and birth weight in Massachusetts, Environ. Health, 11, 40, https://doi.org/10.1186/1476-069X-11-40, 2012.

Kloog, I., Chudnovsky, A. A., Just, A. C., Nordio, F., Koutrakis, P., Coull, B. A., Lyapustin, A., Wang, Y., and Schwartz, J.: A new hybrid spatio-temporal model for estimating daily multi-year PM$_{2.5}$ concentrations across northeastern USA using high resolution aerosol optical depth data, Atmos. Environ., 95, 581–590, 2014.

Koelemeijer, R., Homan, C., and Matthijsen, J.: Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe, Atmos. Environ., 40, 5304–5315, 2006.

Kumar, N., Chu, A., and Foster, A.: An empirical relationship between PM$_{2.5}$ and aerosol optical depth in Delhi Metropolitan, Atmos. Environ., 41, 4492–4503, 2007.

Laden, F., Schwartz, J., Speizer, F. E., and Dockery, D. W.: Reduction in Fine Particulate Air Pollution and Mortality, Am. J. Resp. Crit. Care, 173, 667–672, 2006.

Lee, H. J., Liu, Y., Coull, B. A., Schwartz, J., and Koutrakis, P.: A novel calibration approach of MODIS AOD data to predict PM$_{2.5}$ concentrations, Atmos. Chem. Phys., 11, 7991–8002, https://doi.org/10.5194/acp-11-7991-2011, 2011.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989–3034, https://doi.org/10.5194/amt-6-2989-2013, 2013.

Li, S., Kahn, R., Chin, M., Garay, M. J., and Liu, Y.: Improving satellite-retrieved aerosol microphysical properties using GOCART data, Atmos. Meas. Tech., 8, 1157–1171, https://doi.org/10.5194/amt-8-1157-2015, 2015.

Limbacher, J. A. and Kahn, R. A.: MISR research-aerosol-algorithm refinements for dark water retrievals, Atmos. Meas. Tech., 7, 3989–4007, https://doi.org/10.5194/amt-7-3989-2014, 2014.

Limbacher, J. A. and Kahn, R. A.: MISR empirical stray light corrections in high-contrast scenes, Atmos. Meas. Tech., 8, 2927–2943, https://doi.org/10.5194/amt-8-2927-2015, 2015.

Limbacher, J. A. and Kahn, R. A.: Updated MISR dark water research aerosol retrieval algorithm – Part 1: Coupled

1.1 km ocean surface chlorophyll *a* retrievals with empirical calibration corrections, Atmos. Meas. Tech., 10, 1539–1555, https://doi.org/10.5194/amt-10-1539-2017, 2017.

Liu, Y., Park, R. J., Jacob, D. J., Li, Q., Kilaru, V., and Sarnat, J. A.: Mapping annual mean ground-level PM$_{2.5}$ concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States, J. Geophys. Res., 109, D22206, https://doi.org/10.1029/2004JD005025, 2004.

Liu, Y., Koutrakis, P., and Kahn, R.: Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 1 – Method development, J. Air Waste Manage., 57, 1351–1359, 2007a.

Liu, Y., Koutrakis, P., Kahn, R., Turquety, S., and Yantosca, R. M.: Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 2 – A case study, J. Air Waste Manage., 57, 1360–1369, 2007b.

Lv, B., Hu, Y., Chang, H. H., Russell, A. G., and Bai, Y.: Improving the accuracy of daily PM$_{2.5}$ distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in North China, Environ. Sci. Technol., 50, 4752–4759, 2016.

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., Levy, R., and Reid, J. S.: Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm, J. Geophys. Res., 116, D03211, https://doi.org/10.1029/2010JD014986, 2011.

Lyapustin, A., Wang, Y., Laszlo, I., and Korkin, S.: Improved cloud and snow screening in MAIAC aerosol retrievals using spectral and spatial analysis, Atmos. Meas. Tech., 5, 843–850, https://doi.org/10.5194/amt-5-843-2012, 2012.

Lyapustin, A., Wang, Y., Korkin, S., and Huang, D.: MODIS Collection 6 MAIAC Algorithm, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2018-141, in review, 2018.

Ma, Z., Hu, X., Huang, L., Bi, J., and Liu, Y.: Estimating ground-level PM$_{2.5}$ in China using satellite remote sensing, Environ. Sci. Technol., 48, 7436–7444, 2014.

Ma, Z., Liu, Y., Zhao, Q., Liu, M., Zhou, Y., and Bi, J.: Satellite-derived high resolution PM$_{2.5}$ concentrations in Yangtze River Delta Region of China using improved linear mixed effects model, Atmos. Environ., 133, 156–164, 2016.

Malm, W. C., Sisler, J. F., Huffman, D., Eldred, R. A., and Cahill, T. A.: Spatial and seasonal trends in particle concentration and optical extinction in the United States. J. Geophys. Res.-Atmos., 99, 1347–1370, 1994.

Malm, W. C., Day, D. E., and Kreidenweis, S. M.: Light scattering characteristics of aerosols as a function of relative humidity: Part I – A comparison of measured scattering and aerosol concentrations using the theoretical models, J. Air Waste Manage., 50, 686–700, 2000.

Malm, W. C., Schichtel, B. A., and Pitchford, M. L.: Uncertainties in PM$_{2.5}$ gravimetric and speciation measurements and what we can learn from them, J. Air Waste Manage., 61, 1131–1149, 2011.

Mathur, R.: Estimating the impact of the 2004 Alaskan forest fires on episodic particulate matter pollution over the eastern United States through assimilation of satellite-derived aerosol optical depths in a regional air quality model, J. Geophys. Res., 113, D17302, https://doi.org/10.1029/2007JD009767, 2008.

Monks, P., Granier, C., Fuzzi, S., Stohl, A., Williams, M., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., and Bey, I.: Atmospheric composition change – global and regional air quality, Atmos. Environ., 43, 5268–5350, 2009.

Ngo, M., Pinkerton, K., Freeland, S., Geller, M., Ham, W., Cliff, S., Hopkins, L., Kleeman, M., Kodavanti, U., and Meharg, E.: Airborne particles in the San Joaquin Valley may affect human health, Calif. Agr., 64, 12–16, 2010.

Nolte, C. G., Appel, K. W., Kelly, J. T., Bhave, P. V., Fahey, K. M., Collett Jr., J. L., Zhang, L., and Young, J. O.: Evaluation of the Community Multiscale Air Quality (CMAQ) model v5.0 against size-resolved measurements of inorganic particle composition across sites in North America, Geosci. Model Dev., 8, 2877–2892, https://doi.org/10.5194/gmd-8-2877-2015, 2015

Özkaynak, H., Frey, H. C., Burke, J., and Pinder, R. W.: Analysis of coupled model uncertainties in source-to-dose modeling of human exposures to ambient air pollution: A PM$_{2.5}$ case study, Atmos. Environ., 43, 1641–1649, 2009.

Park, R. S., Song, C. H., Han, K. M., Park, M. E., Lee, S.-S., Kim, S.-B., and Shimizu, A.: A study on the aerosol optical properties over East Asia using a combination of CMAQ-simulated aerosol optical properties and remote-sensing data via a data assimilation technique, Atmos. Chem. Phys., 11, 12275–12296, https://doi.org/10.5194/acp-11-12275-2011, 2011.

Patadia, F., Kahn, R. A., Limbacher, J. A., Burton, S. P., Ferrare, R. A., Hostetler, C. A., and Hair, J. W.: Aerosol airmass type mapping over the Urban Mexico City region from space-based multi-angle imaging, Atmos. Chem. Phys., 13, 9525–9541, https://doi.org/10.5194/acp-13-9525-2013, 2013.

Pinto, J. P., Lefohn, A. S., and Shadwick, D. S.: Spatial variability of PM$_{2.5}$ in urban areas in the United States, J. Air Waste Manage., 54, 440–449, 2004.

Pope, C. A., Ezzati, M., and Dockery, D. W.: Fine-Particulate Air Pollution and Life Expectancy in the United States, New Engl. J. Med., 360, 376–386, 2009.

Rees, S. L., Robinson, A. L., Khlystov, A., Stanier, C. O., and Pandis, S. N.: Mass balance closure and the Federal Reference Method for PM$_{2.5}$ in Pittsburgh, Pennsylvania, Atmos. Environ., 38, 3305–3318, 2004.

Sarwar, G., Simon, H., Bhave, P., and Yarwood, G.: Examining the impact of heterogeneous nitryl chloride production on air quality across the United States, Atmos. Chem. Phys., 12, 6455–6473, https://doi.org/10.5194/acp-12-6455-2012, 2012.

Schaap, M., Apituley, A., Timmermans, R. M. A., Koelemeijer, R. B. A., and de Leeuw, G.: Exploring the relation between aerosol optical depth and PM$_{2.5}$ at Cabauw, the Netherlands, Atmos. Chem. Phys., 9, 909–925, https://doi.org/10.5194/acp-9-909-2009, 2009.

Schoups, G., Hopmans, J. W., Young, C. A., Vrugt, J. A., Wallender, W. W., Tanji, K. K., and Panday, S.: Sustainability of irrigated agriculture in the San Joaquin Valley, California, P. Natl. Acad. Sci. USA, 102, 15352–15356, 2005.

Shi, Y., Zhang, J., Reid, J. S., Liu, B., and Hyer, E. J.: Critical evaluation of cloud contamination in the MISR aerosol products using MODIS cloud mask products, Atmos. Meas. Tech., 7, 1791–1801, https://doi.org/10.5194/amt-7-1791-2014, 2014.

Simon, H., Bhave, P. V., Swall, J. L., Frank, N. H., and Malm, W. C.: Determining the spatial and seasonal variability in OM / OC ratios across the US using multiple regression, Atmos. Chem. Phys., 11, 2933–2949, https://doi.org/10.5194/acp-11-2933-2011, 2011.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF Version 3, NCAR technical note, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA, 2008.

Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., and Slutsker, I.: Cloud-screening and quality control algorithms for the AERONET database, Remote Sens. Environ., 73, 337–349, 2000.

Solomon, P. A., Crumpler, D., Flanagan, J. B., Jayanty, R. K. M., Rickman, E. E., and McDade, C. E.: U.S. National PM$_{2.5}$ Chemical Speciation Monitoring Networks – CSN and IMPROVE: Description of networks, J. Air Waste Manage., 64, 1410–1438, 2014.

Song, C. H., Park, M. E., Lee, K. H., Ahn, H. J., Lee, Y., Kim, J. Y., Han, K. M., Kim, J., Ghim, Y. S., and Kim, Y. J.: An investigation into seasonal and regional aerosol characteristics in East Asia using model-predicted and remotely-sensed aerosol properties, Atmos. Chem. Phys., 8, 6627–6654, https://doi.org/10.5194/acp-8-6627-2008, 2008.

Song, W., Jia, H., Huang, J., and Zhang, Y.: A satellite-based geographically weighted regression model for regional PM$_{2.5}$ estimation over the Pearl River Delta region in China, Remote Sens. Environ., 154, 1–7, 2014.

Tanner, R. L., Parkhurst, W. J., Valente, M. L., and Phillips, W. D.: Regional composition of PM$_{2.5}$ aerosols measured at urban, rural and "background" sites in the Tennessee valley, Atmos. Environ., 38, 3143–3153, 2004.

Tong, D. Q. and Mauzerall, D. L.: Spatial variability of summertime tropospheric ozone over the continental United States: Implications of an evaluation of the CMAQ model, Atmos. Environ., 40, 3041–3056, 2006.

Tsai, T.-C., Jeng, Y.-J., Chu, D. A., Chen, J.-P., and Chang, S.-C.: Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008, Atmos. Environ., 45, 4777–4788, 2011.

US EPA – US Environmental Protection Agency: Air Quality Criteria for Particulate Matter (Final Report), EPA 600/P-99/002aF-bF, Washington, D.C., 2004.

US EPA – US Environmental Protection Agency: 2011 National Emissions Inventory Technical Support Document, version 2, https://www.epa.gov/sites/production/files/2015-10/documents/nei2011v2_tsd_14aug2015.pdf (last access: 7 August 2018), 2015.

US EPA – US Environmental Protection Agency: The Green Book Nonattainment Areas for Criteria Pollutants, http://www.epa.gov/green-book, last access: 7 August 2018.

Van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P. J.: Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application, Environ. Health Persp., 118, 847–855, https://doi.org/10.1289/ehp.0901623, 2010.

van Donkelaar, A., Martin, R. V., Spurr, R. J., Drury, E., Remer, L. A., Levy, R. C., and Wang, J.: Optimal estimation for global ground-level fine particulate matter concentrations, J. Geophys. Res.-Atmos., 118, 5621–5636, 2013.

Wade, K. S., Mulholland, J. A., Marmur, A., Russell, A. G., Hartsell, B., Edgerton, E., Klein, M., Waller, L., Peel, J. L., and Tolbert, P. E.: Effects of instrument precision and spatial variability on the assessment of the temporal variation of ambient air pollution in Atlanta, Georgia, J. Air Waste Manage., 56, 876–888, 2006.

Wallace, J. and Kanaroglou, P.: An investigation of air pollution in southern Ontario, Canada, with MODIS and MISR aerosol data, in: 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, 4311–4314, 2007.

Wang, J. and Christopher, S. A.: Intercomparison between satellite-derived aerosol optical thickness and PM$_{2.5}$ mass: implications for air quality studies, Geophys. Res. Lett., 30, 2095, https://doi.org/10.1029/2003GL018174, 2003.

Wang, J., Xu, X., Spurr, R., Wang, Y., and Drury, E.: Improved algorithm for MODIS satellite retrievals of aerosol optical thickness over land in dusty atmosphere: Implications for air quality monitoring in China, Remote Sens. Environ., 114, 2575–2583, 2010.

Wang, J., Aegerter, C., Xu, X., and Szykman, J. J.: Potential application of VIIRS Day/Night Band for monitoring nighttime surface PM$_{2.5}$ air quality from space, Atmos. Environ., 124, 55–63, 2016.

Watson, J. G. and Chow, J. C.: A wintertime PM$_{2.5}$ episode at the Fresno, CA, supersite, Atmos. Environ., 36, 465–475, 2002.

Whitten, G. Z., Heo, G., Kimura, Y., McDonald-Buller, E., Allen, D. T., Carter, W. P., and Yarwood, G.: A new condensed toluene mechanism for Carbon Bond: CB05-TU, Atmos. Environ., 44, 5346–5355, 2010.

Witek, M. L., Garay, M. J., Diner, D. J., and Smirnov, A.: Aerosol optical depths over oceans: A view from MISR retrievals and collocated MAN and AERONET in situ observations, J. Geophys. Res.-Atmos., 118, 12620–12633, https://doi.org/10.1002/2013JD020393, 2013.

Yarwood, G., Rao, S., Yocke, M., and Whitten, G.: Updates to the carbon bond chemical mechanism: CB05, Final report to the US EPA, RT-0400675, 8, available at: http://www.camx.com/publ/pdfs/CB05_Final_Report_120805.pdf (last access: 7 August 2018), 2005.

Zhang, H., Hoff, R. M., and Engel-Cox, J. A.: The relation between Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth and PM$_{2.5}$ over the United States: a geographical comparison by US Environmental Protection Agency regions, J. Air Waste Manage., 59, 1358–1369, 2009.

Remarks from the language copy-editor

Remarks from the typesetter